

- 7 APR. 2003

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

(31)

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 1/19, 1/15, 15/80, 15/81	A1	(11) International Publication Number: WO 98/12300 (43) International Publication Date: 26 March 1998 (26.03.98)
---	-----------	--

(21) International Application Number: PCT/DK97/00397

(22) International Filing Date: 19 September 1997 (19.09.97)

(30) Priority Data:
1024/96 19 September 1996 (19.09.96) DK(71) Applicant (for all designated States except US): NOVO
NORDISK A/S [DK/DK]; Novo Allé, DK-2880 Bagsværd (DK).

(72) Inventor; and

(75) Inventor/Applicant (for US only): LEHMBECK, Jan
[DK/DK]; Novo Nordisk a/s, Novo Allé, DK-2880
Bagsværd (DK).(74) Common Representative: NOVO NORDISK A/S; Corporate
Patents, Novo Allé, DK-2880 Bagsværd (DK).(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR,
BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE,
GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO,
NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,
TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH,
KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ,
BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE,
CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,
PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN,
ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: NOVEL HOST CELLS AND METHODS OF PRODUCING PROTEINS

(57) Abstract

The present invention relates to novel host cells and to methods of producing proteins. More specifically the invention relates to a host cell useful for the expression of heterologous proteins, in which the host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease and an alkaline protease. Moreover the invention relates to a method of producing a heterologous protein, which method comprises cultivating the host cell in a suitable growth medium, followed by recovery of the desired protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

NOVEL HOST CELLS AND METHODS OF PRODUCING PROTEINS

FIELD OF THE INVENTION

5

The present invention relates to novel fungal host cells and to methods of producing proteins. More specifically, the invention relates to a host cell useful for the expression of heterologous proteins in which the host cell has been genetically modified in order to express significantly reduced levels of a metalloprotease and an alkaline protease. Moreover, the
10 invention relates to a method of producing proteins of interest in high yields by using the fungi of the invention, in which the method comprises cultivating the host cell in a suitable growth medium, followed by recovery of the desired protein. The invention also comprises methods for producing such fungi and DNA constructs to be used in these methods.

15

BACKGROUND OF THE INVENTION

The use of recombinant host cells in the expression of heterologous proteins has in recent years greatly simplified the production of large quantities of commercially valuable proteins
20 which otherwise are obtainable only by purification from their native sources. Currently, there is a varied selection of expression systems from which to choose for the production of any given protein, including eubacterial and eukaryotic hosts. The selection of an appropriate expression system often depends not only on the ability of the host cell to produce adequate yields of the protein in an active state, but, to a large extent, may also be governed by the intended end use of
25 the protein.

One problem frequently encountered is the high level of proteolytic enzymes produced by a given host cell or present in the culture medium. It has been suggested that one could provide host organisms deprived of the ability to produce specific proteolytic compounds. For example, International Patent Application WO 90/00192 (Genencor) describes filamentous fungal hosts
30 incapable of secreting enzymatically active aspartic proteinase, and EP 574 347 (Ciba Geigy AG) describes *Aspergillus* hosts defective in a serine protease of the subtilisin-type.

1 Metalloproteases have been isolated from a number of eukaryotic sources. Neutral metalloproteases, i.e. metalloproteases having optimal activity at neutral pH, isolated from strains of *Aspergillus* also have been reported. The physicochemical properties of two neutral metalloproteases from *Aspergillus sojae*, NpI and NpII, have been reported in a series of studies
5 (Sekine, H. 1972. Agric. Biol. Chem. 36:198-206, 207-216; Sekine, H. 1972. Agric. Biol. Chem. 36:2143-2150). It was revealed that the enzymatic and physicochemical characteristics of NpI, but not Np II, resemble those of *Bacillus thermoproteolyticus* thermolysin. Recently, the cDNA sequence of a neutral metalloprotease, NpII, from *Aspergillus oryzae* was disclosed (Tatsumi H, et al. 1991. Mol. Gen. Genet. 228:97-103). However, the cDNA sequence of a
10 neutral metalloprotease from the NpI group from *Aspergillus oryzae* has never been disclosed.

Alkaline protease is a serine protease with an alkaline pH optimum (Nakagawa, Y. 1970. Methods Enzymol. 19:581-591). It is a homologue of subtilisin and is the predominant extracellular alkaline endopeptidase of *Aspergillus oryzae*. The gene has been isolated and characterised (Murakami, et al. 1991. Agric. Biol. Chem. 55:2807-2811). Two other species of
15 *Aspergillus*, *A. flavus* and *A. sojae*, produce identical or closely related enzymes.

A potential role of metalloproteases and alkaline proteases in reducing the stability of protein products obtained from *Aspergillus* has not been reported.

20 SUMMARY OF THE INVENTION

According to the present invention it has now been found that the proteolytic activity of neutral metalloprotease I and alkaline protease, either individually or in combination, may significantly reduce the stability of a protein product of interest produced by a cell resulting in a
25 reduced yield of the said protein.

Accordingly, the present invention provides a host cell useful for the expression of a heterologous protein product, in which the cell has been genetically modified in order to express significantly reduced levels of both a metalloprotease and an alkaline protease, by comparison to the parental cell.

30 In another aspect, the invention provides a method of producing a heterologous protein product in a host cell of the invention, in which the method comprises introducing into the host

cell a nucleic acid sequence encoding the protein, cultivating the host cell in a suitable growth medium, and recovering the heterologous protein product.

By the method of the invention, proteolytic activity resulting from a neutral metalloprotease I and an alkaline protease are significantly reduced, thereby improving the stability and yield of the protein obtained by the method. Moreover, the protein obtained by the method of the invention may be a precursor protein such as a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or any other type of immature form.

10 BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is further illustrated by reference to the accompanying drawings, in which:

- Fig. 1 diagrams the construction of the plasmid pJaL335 which harbours the *pyrG* gene;
- 15 Fig. 2 diagrams the construction of the plasmid pSO5 which harbours 5' and 3' sequences from *pyrG* gene;
- Fig. 3 diagrams the construction of the plasmid pJaL212 in which the *pyrG* coding sequence has been inserted between 5' and 3' sequences of the *alp* gene;
- Fig. 4 diagrams the construction of the plasmid pJaL399 in which the NpI encoding
- 20 sequence is disrupted;
- Fig. 5 shows the N-terminal amino acid sequence of *Candida antarctica* lipase B and the sequences of two oligonucleotide primers derived from this sequence;
- Fig. 6 diagrams the construction of plasmid pMT1305 which harbours the *C. antarctica* lipase B gene;
- 25 Fig. 7 diagrams the construction of the plasmid pMT1329 which harbours partial sequence from the 5' end of the *C. antarctica* lipase B gene;
- Fig. 8 diagrams the construction of the plasmid pMT1332 which harbours the *C. antarctica* lipase B gene; and
- 30 Fig. 9 diagrams the construction of the expression plasmid pMT1335 which harbours the *C. antarctica* lipase B gene.

DETAILED DISCLOSURE OF THE INVENTION

Host Cells

5 The present invention provides a host cell useful for the expression of heterologous proteins, wherein the cell has been genetically modified in order to express significantly reduced levels of metalloprotease and alkaline protease activity in comparison to a parental cell. Said host cell is derived from the parental cell which may be a wild type cell.

The host cell of the invention may be any host cell conventionally used for the heterologous
10 expression of proteins.

Preferably, the host cell of the invention is a yeast or a filamentous fungus capable of producing a desired protein. In particular, the yeast cell may be a strain of *Saccharomyces*, preferably *Saccharomyces cerevisiae*. In particular, the filamentous fungus may be a strain selected from the group consisting of *Acremonium*, *Aspergillus*, *Candida*, *Cocliobolus*, *Endothia*,
15 *Fusarium*, *Humicola*, *Neurospora*, *Rhizomucor*, *Rhizopus*, *Thermomyces*, *Trichoderma*, *Podospora*, *Pyricularia*, and *Penicillium*.

In a preferred embodiment, the filamentous fungus is a strain selected from the group consisting of *Aspergillus nidulans*, *Aspergillus awamori*, *Aspergillus phoenicis*, *Aspergillus japonicus*, *Aspergillus foetus*, *Fusarium graminearum*, *Fusarium oxysporum*, *Fusarium solani*,
20 *Fusarium venenatum*, *Humicola grisea*, *Neurospora crassa*, *Penicillium chrysogenum*, *Rhizomucor meihei*, *Trichoderma reesei*, and *Trichoderma viride*; in particular, a strain of *Aspergillus oryzae*, *Aspergillus niger* and *Fusarium* species having the identifying characteristics of ATCC 20334.

25 Metalloproteases

In the context of this invention, a metalloprotease is a proteolytic enzyme containing a catalytic zinc metal center which participates in the hydrolysis of the peptide backbone of the substrate. The active zinc centre differentiates these proteases from the calpains, whose activities are dependent upon the presence of calcium. Confirmation of a protease as a metalloprotease is
30 by the reversible loss of proteolytic activity upon removal of the zinc centre with 1 mM 1,10-phenanthroline, and its restoration after titration with Zn^{2+} (0.1-100 mM).

In a preferred embodiment, the metalloprotease contemplated in the context of this invention is a *Fusarium* metalloprotease, preferably a *Fusarium oxysporum* metalloprotease. In a most preferred embodiment, the metalloprotease is a *Fusarium oxysporum* p45 metalloprotease (p45) having the nucleotide sequence presented as SEQ. ID. No. 1, or a sequence homologous thereto.

In another preferred embodiment, the metalloprotease contemplated in the context of this invention is a neutral metalloprotease, which is a metalloprotease possessing optimal proteolytic activity in the neutral pH region, i.e. in the range of about pH 6-8, preferably in the range of about pH 6.5-7.5, more specifically, around pH 7. More particularly, the metalloprotease contemplated in the context of this invention is a neutral *Aspergillus* metalloprotease of the group NpI or NpII (Tatsumi, et al., 1991, *supra*).

In a preferred embodiment, the metalloprotease is an *Aspergillus oryzae* neutral Metalloprotease I (NpI) comprising an amino acid sequence derived from the nucleotide sequence presented as SEQ. ID. No. 2, or a sequence homologous thereto.

Alkaline proteases

In the context of this invention an alkaline protease is a serine protease with activity which peaks in the neutral to alkaline pH range. Analyses of the amino acid sequence of alkaline proteases indicate homology to the subtilase subgroup of subtilisin-like serine proteases. As summarised by Siezen, et al. (1991. Protein Eng. 4:719-737) more than 50 subtilases have been identified from a wide variety of organisms, ranging from various species of bacteria, including gram positive and gram negative species, to fungi and yeast to higher eukaryotes, including worms, insects, plants and mammals. The amino acid sequence has been determined in more than 40 of these subtilases, and reveal that the mature region of the enzyme ranges from 268 to 1775 amino acids in length and a pre-pro-region of 27 to 280 amino acids in the N-terminal vicinity. In fungi and yeast, the variation is apparently smaller, with corresponding ranges of 279 to 281 and 105 to 121 in fungi, and 297 to 677 and 126 to 280 in yeast. A cDNA fragment of the entire coding region of the alkaline protease from *Aspergillus oryzae* was cloned and expressed in *Saccharomyces cerevisiae* (Tatsumi, H, et al. 1989. Mol. Gen. Genet. 219:33-38). The primary structure was shown to share 29% to 44% homology with other sequenced subtilisins, and the three residues in the active site, Asp32, His 64 and Ser221 in subtilisin BPN', were conserved.

In a preferred embodiment, the alkaline protease is an *Aspergillus oryzae* alkaline protease (*alp*), preferably encoded by a cDNA sequence comprising the nucleotide sequence presented as SEQ. ID. No. 3, or a sequence homologous thereto.

5 Sequence Homology

As used herein DNA sequence homology is determined as the degree of identity between the two sequences indicating a derivation of the first sequence from the second. The homology may be suitably determined by means of computer programs well known in the art, such as GAP, provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, 10 August 1994, Genetics Computer Group, 575 Science Drive, Madison WI, USA; Needleman, S.B. and Wunsch, C.D., 1970. J. Mol. Biol. 48:443-453). Using the following GAP settings for DNA sequence comparison, a creation penalty of 5.0 and an extension penalty of 0.3, the coding region of the analogous DNA sequence exhibits a degree of identity preferably of at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, 15 and most preferably at least 97% with the encoding region of an indicated DNA sequence.

Amino acid sequence homology as used herein is similarly determined as the degree of identity between two sequences indicating a derivation of the first sequence from the second. The homology may be suitably determined by means of computer programs well known in the art, such as GAP referred to above (Needleman, S.B. and Wunsch, C.D., 1970, *supra*). 20 Using GAP with the following settings for amino acid sequence comparison, a creation penalty of 3.0 and extension penalty of 0.1, the polypeptide encoded by an analogous DNA sequence exhibits a degree of identity preferably of at least 70%, more preferably at least 80%, more preferably at least 90%, more preferably at least 95%, especially at least 97% with a DNA sequence comprising the region encoding the structural protein.

25 The present invention is also directed to variants of metalloproteases and alkaline proteases which have an amino acid sequence differing by no more than three amino acids, preferably by no more than two amino acids, and more preferably by no more than one amino acid from the mature polypeptide of the amino acid sequence derived from SEQ. ID. No. 1, SEQ. ID No. 2 and SEQ. ID. No. 3.

30 Hybridisation as used herein is intended to indicate that an analogous DNA sequence binds to an oligonucleotide probe corresponding to a polypeptide encoded by an indicated

DNA sequence comprising the region encoding the structural protein under certain specified conditions which are described in detail below.

Suitable conditions for determining hybridization between a nucleotide probe and a homologous DNA or RNA sequence involves presoaking the filter containing the DNA fragments or RNA to be hybridized in 5x SSC (standard saline citrate buffer; *cf*, Sambrook et al., 1989. Molecular Cloning, Cold Spring Harbor NY, USA) for 10 min, and prehybridization of the filter in a solution of 5x SSC, 5 x Denhardt's solution, 0.5 % SDS and 100 µg/ml of denatured sonicated salmon sperm DNA (Sambrook et al. 1989, *supra*), followed by hybridization in the same solution containing a random-primed (Feinberg, A. P., and Vogelstein, B. 1983. Anal. Biochem. 132:6-13), ³²P-dCTP-labeled (specific activity > 1 x 10⁹ cpm/µg) probe for 12 hours at ca. 45°C. The filter is then washed two times for 30 minutes in 2x SSC, 0.5 % SDS at a temperature of at least 55°C (low stringency), more preferably at least 60°C (medium stringency), more preferably at least 65°C (medium/high stringency), more preferably at least 70°C (high stringency), even more preferably at least 75°C (very high stringency). Molecules to which the oligonucleotide probe hybridise under these conditions are detected by exposure to x-ray film.

Genetic Modifications of the Host Cell

The host cell of the invention, in order to express significantly reduced levels of metalloprotease and alkaline protease activity, is genetically modified which may be achieved by using standard recombinant DNA technology known to the person skilled in the art. The gene sequences respectively responsible for production of metalloprotease and alkaline protease activity may be inactivated or partially or entirely eliminated. Thus, a host cell of the invention expresses reduced or undetectable levels of metalloprotease and alkaline protease or expresses functionally inactive metalloprotease and alkaline protease.

In a particular embodiment, the said inactivation is obtained by modification of the respective structural or regulatory regions encoded within the metalloprotease and alkaline protease genes of interest. Known and useful techniques include, but are not limited to, specific or random mutagenesis, PCR generated mutagenesis, site specific DNA deletion, insertion and/or substitution, gene disruption or gene replacement, anti-sense techniques, or a combination thereof.

Mutagenesis may be performed using a suitable physical or chemical mutagenising agent. Examples of a physical or chemical mutagenising agent suitable for the present purpose include, but are not limited to, ultraviolet (UV) irradiation, hydroxylamine, N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), O-methyl hydroxylamine, nitrous acid, ethyl methane sulphonate
5 (EMS), sodium bisulfite, formic acid, and nucleotide analogues. When such agents are used, the mutagenesis is typically performed by incubating the cell to be mutagenised in the presence of the mutagenising agent of choice under suitable conditions, and selecting for cells showing a significantly reduced production of NpI and alp.

Modification may also be accomplished by the introduction, substitution or removal of one
10 or more nucleotides in the structural sequence or a regulatory element required for the transcription or translation of the structural sequence. For example, nucleotides may be inserted or removed so as to result in the introduction of a stop codon, the removal of the start codon or a change of the open reading frame of the structural sequence. The modification or inactivation of the structural sequence or a regulatory element thereof may be accomplished by site-directed
15 mutagenesis or PCR generated mutagenesis in accordance with methods known in the art. Although, in principle, the modification may be performed *in vivo*, i.e. directly on the cell expressing the metalloprotease and alkaline protease genes, it is presently preferred that the modification be performed *in vitro* as exemplified below.

A convenient way to inactivate or reduce the metalloprotease and alkaline protease
20 production in a host cell of choice is based on techniques of gene interruption. In this method a DNA sequence corresponding to the endogenous gene or gene fragment of interest is mutagenised *in vitro*. Said DNA sequence thus encodes a defective gene which is then transformed into the host cell. By homologous recombination, the defective gene replaces the endogenous gene or gene fragment. It may be desirable that the defective gene or gene fragment also encodes a
25 marker which may be used to select for transformants in which the respective genes encoding metalloprotease and/or alkaline protease have been modified or destroyed.

Methods for deleting or disrupting a targeted gene are specifically described by Miller, et al (1985. Mol. Cell. Biol. 5:1714-1721); WO 90/00192; May, G. (1992. Applied Molecular Genetics of Filamentous Fungi. J.R. Kinghorn and G. Turner, eds., Blackie Academic and
30 Professional, pp. 1-25); and Turner, G. (1994. Vectors for Genetic Manipulation. S.D. Martinelli and J.R. Kinghorn, eds., Elsevier, pp.641-665).

Alternatively, the modification or inactivation of the DNA sequence may be performed by established anti-sense techniques using a nucleotide sequence complementary to an encoding sequence for metalloprotease, e.g. the nucleotide sequences presented as SEQ. ID. No. 1 and SEQ. ID. No. 2, or an alkaline protease encoding sequence, e.g. the nucleotide sequence shown
5 in SEQ. ID. No. 4. The anti-sense technology and its application is described in detail in US Patent No. 5,190,931 (University of New York).

Therefore, due to genetic modification, the host cell of the invention expresses significantly reduced levels of metalloprotease and alkaline protease activity. In a preferred embodiment, the level of these proteolytic activities expressed by the host cell is individually reduced more than
10 about 50%, preferably more than about 85%, more preferably more than about 90%, and most preferably more than about 95%. In another preferred embodiment, the these proteolytic activities in the host cell of the invention may be reduced in any combination. In a most preferred embodiment, the product expressed by the host cell is essentially free from proteolytic activity due to metalloprotease and alkaline protease.

15

Methods of Producing Proteins

By the method of the invention, the proteolytic activities of metalloprotease and alkaline protease are significantly reduced, thereby improving the stability and increasing the yield of susceptible protein products synthesised by the host cell of the invention. More specifically, by
20 the method of the invention, the host cell is genetically modified within structural and/or regulatory regions encoded within the metalloprotease and alkaline protease genes.

Therefore, another aspect of the invention provides a method of producing proteins in a host cell of the invention, including heterologous polypeptides, in which the method comprises introducing into said host cell a nucleic acid sequence encoding the protein product of interest,
25 cultivating the host cell in a suitable growth medium, followed by recovery of the protein product.

Thus, the host cell of the invention must contain structural and regulatory genetic regions necessary for the expression of the desired product. The nature of such structural and regulatory regions greatly depends on the product and the host cell in question. The genetic design of the
30 host cell of the invention may be accomplished by the person skilled in the art using standard recombinant DNA technology for the transformation or transfection of a host cell (*vide, e.g., Sambrook et al., inter alia*).

Preferably, the host cell is modified by methods known in the art for the introduction of an appropriate cloning vehicle, i.e. a plasmid or a vector, comprising a DNA fragment encoding the desired protein product. The cloning vehicle may be introduced into the host cell either as an autonomously replicating plasmid or integrated into the chromosome. Preferably, the cloning
5 vehicle comprises one or more structural regions operably linked to one or more appropriate regulatory regions.

The structural regions are regions of nucleotide sequences encoding the desired protein product. The regulatory regions include promoter regions comprising transcription and translation control sequences, terminator regions comprising stop signals, and polyadenylation
10 regions. The promoter, i.e. a nucleotide sequence exhibiting a transcriptional activity in the host cell of choice, may be one derived from a gene encoding an extracellular or an intracellular protein, preferably an enzyme, such as an amylase, a glucoamylase, a protease, a lipase, a cellulase, a xylanase, an oxidoreductase, a pectinase, a cutinase, or a glycolytic enzyme. Examples of suitable promoters for transcription in a fungal host cell are promoters derived from
15 the gene encoding *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* neutral α -amylase, *Aspergillus niger* acid stable α -amylase, *Aspergillus niger* or *Aspergillus awamsii* glucoamylase (GluA), *Aspergillus nidulans* acetamidase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphatase isomerase, *Rhizopus meihei* aspartic proteinase, and *Rhizopus meihei* lipase. Preferred promoters are the *Aspergillus oryzae* TAKA-amylase and *Aspergillus awamsii*
20 GluA.

The cloning vehicle may also include a selectable marker, such as a gene product which complements a defect in the host cell, or one which confers antibiotic resistance. Examples of antibiotics useful as *Aspergillus* selection markers include hygromycin, phleomycin and basta. Other examples of *Aspergillus* selection markers include *amdS*, which encodes an enzyme
25 involved in acetamide utilisation; *pyrG*, which encodes an enzyme involved in uridine biosynthesis; *argB*, which encodes an enzyme involved in arginine biosynthesis; *niaD*, which encodes an enzyme involved in the nitrate assimilation pathway; and *sC*, which encodes an enzyme involved in the sulfate assimilation pathway. Preferred for use in an *Aspergillus* host cell are the *amdS* and *pyrG* markers of *Aspergillus nidulans* or *Aspergillus oryzae*. Furthermore,
30 selection may be accomplished by co-transformation, wherein the transformation is carried out with a mixture of two vectors and the selection is made for one vector only.

The procedures used to ligate the DNA construct of the invention, the promoter, terminator and other elements, respectively, and to insert them into suitable cloning vehicles containing the information necessary for replication, are well known to persons skilled in the art (*vide e.g.*, Sambrook et al., 1989; *inter alia*).

5 The culture broth or medium used may be any conventional medium suitable for culturing the host cell of the invention, and formulated according to the principles of the prior art. The medium preferably contains carbon and nitrogen sources as well as other inorganic salts. Suitable media, e.g. minimal or complex media, are available from commercial suppliers, or may be prepared according to published recipes, as in The Catalogue of Strains, published by The
10 American Type Culture Collection. Rockville MD, USA. 1970.

The appropriate pH for fermentation will often be dependent on such factors as the nature of the host cell to be used, the composition of the growth medium, the stability of the polypeptide of interest, and the like. Consequently, although the host cell of the invention may be cultured using any fermentation process performed at any pH, it is advantageous that the pH of the
15 fermentation process is such that acidic and/or neutral protease activities of the host cell are essentially eliminated or at least significantly reduced. Thus, removal of aspartic protease activity as described in WO 90/00192 may also be accomplished by raising the fermentation pH, and does not present any additional advantageous effect on the yield of a desired protein from host cells cultivated in the alkaline pH range.

20 If the pH of the fermentation process is within the range from 5 to 11, such as from 6 to 10.5, 7 to 10, or 8 to 9.5, the activity of acidic proteases, such as aspartic and serine proteases, and neutral proteases in the pH ranges above 7, will be reduced or blocked. Examples of enzymes produced under alkaline fermentation conditions include endoglucanases, phytases and protein disulfide isomerases.

25 However, the alkaline pH range will support alkaline protease activity in an unmodified host cell, which, in turn, may potentially result in degradation of the polypeptide product of interest. Consequently, in such cases the inactivation of the gene encoding alkaline protease is especially advantageous.

Inactivation of the alkaline protease gene of the invention is also especially advantageous for
30 certain host cells, as the levels of acidic, neutral and alkaline protease activities vary from species to species. For example, the level of alkaline protease activity in the *Aspergillus oryzae* is higher than in *Aspergillus niger*.

After cultivation, the desired protein is recovered by conventional methods of protein isolation and purification from a culture broth. Well established purification procedures include separating the cells from the medium by centrifugation or filtration, precipitating proteinaceous components of the medium by means of a salt such as ammonium sulphate, and chromatographic methods such as ion exchange chromatography, gel filtration chromatography, affinity chromatography, and the like.

Products

The desired end product, i.e. the heterologous protein expressed by the host cell of the invention, may be any eubacterial or eukaryotic protein or peptide.

As defined herein, a "heterologous protein" is a protein or polypeptide gene product which is not native to the host cell, or is a native protein in which modifications have been made to alter the native sequence, or is a native protein whose expression is quantitatively altered as a result of a manipulation of a native regulatory sequence required for the expression of the native protein, such as a promoter, a ribosome binding site, etc., or other manipulation of the host cell by recombinant DNA techniques.

In a more specific embodiment, the product is a therapeutically active peptide or protein, such as a hormone, in particular insulin, growth hormone, glucagon, or somatostatin; an interleukin, in particular interferon; a haematopoietic growth factor, in particular PDGF (platelet derived growth factor), EPO (erythropoietin), or TPO (thrombopoietin); a protease, in particular factor VII, factor VIII, urokinase, chymosin, or tissue plasminogen activator; or serum albumin.

In another preferred embodiment, the product is an enzyme of fungal or of bacterial origin.

The enzyme is preferably a glycosidase enzyme, e.g. an amylase, in particular an α -amylase (EC 3.2.1.1) or a β -amylase (EC 3.2.1.2); a glucan 1,4- α -glucosidase (EC 3.2.1.3), a cellulase, in particular an endo-1,4- β -glucanase (EC 3.2.1.4) or an endo-1,3(4)- β -glucanase (EC 3.2.1.6); a xylanase, in particular an endo-1,4- β -xylanase (EC 3.2.1.8) or a xylan-endo-1,3- β -xylosidase (EC 3.2.1.32); a polygalacturonase (EC 3.2.1.15); a glucosidase, in particular an α -glucosidase (EC 3.2.1.20) or a β -glucosidase (EC 3.2.1.21); a galactosidase, in particular an α -galactosidase (EC 3.2.1.22) or a β -galactosidase (EC 3.2.1.23); an endoglucanase, in particular an endo-1,3- β -glucanase (EC 3.2.1.39), an endo-1,3- α -glucanase (EC 3.2.1.59), an endo-1,2- β -glucanase (EC 3.2.1.71), or an endo-1,6- β -glucanase (EC 3.2.1.75); and a cellulose-1,4- β -cellobiosidase (EC 3.2.1.91).

In another preferred embodiment the enzyme is a lipolytic enzyme, in particular a lipase, an esterase, a phospholipase, or a lyso-phospholipase.

In another preferred embodiment the enzyme is a phytase, in particular a 3-phytase (EC 3.1.3.8) or a 6-phytase (EC 3.1.3.26).

5 In another preferred embodiment the enzyme is a proteolytic enzyme.

In another preferred embodiment the enzyme is a laccase, a pectinase, or a cutinase, or an oxidoreductase, such as a peroxidase.

In another preferred embodiment the product is a hybrid polypeptide, preferably prochymosin and pro-trypsin-like proteases.

10 Due to the absence of metalloprotease and alkaline protease activities, the heterologous protein expressed by the host cell may also be a precursor protein such as a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or any other immature form.

15 EXAMPLES

The invention is further illustrated with reference to the following examples which should not in any way be construed as limiting the scope of the invention as defined in the appended claims.

20

Materials and Methods

Strains

25 *Aspergillus oryzae* IFO 4177: available from Institute for Fermentation, Osaka; 17-25 Juso Hammachi 2-Chome Yodogawaku, Osaka, Japan.

30 *Fusarium oxysporum* DSM 2672: deposited according to the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure at Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Mascheroder Weg 1b, DE-3300 Braunschweig, Germany, on 6 June 1983.

HowB101: The construction of this strain is described in Example 1.

35 JaL125: The construction of this strain is described in Example 1.

JaL151: The construction of this strain is described in Example 1.

JaL228: The construction of this strain is described in Example 1.

Genes

alp: This gene codes for the alkaline protease shown in SEQ. ID. No. 3.

Npl: This gene codes for the neutral metalloprotease I shown in SEQ. ID. No. 2.

5 *pyrG*: This gene codes for orotidine-5'-phosphate decarboxylase, an enzyme involved in the biosynthesis of uridine.

p45: This gene codes for the neutral metalloprotease I shown in SEQ. ID. No. 1.

10

Plasmids

pSO2: The preparation is described in Example 1.

pSO5: The preparation is described in Example 1.

15

pSRe403: The preparation is described in Example 1.

PSRe403ΔSacII: The preparation is described in Example 1.

20 *pJaL173*: The preparation is described in Example 1.

pJaL212: The preparation is described in Example 1.

pJaL335: The preparation is described in Example 1.

25

pJaL389: The preparation is described in Example 1.

pJaL399: The preparation is described in Example 1.

30 *pToC65*: The construction is described in EP 531 372 B.

pToC68: The construction is described in WO 91/17243.

35 *pToC90*: A subclone of *p3SR2* (Corrick et al., 1987. GENE 53:63-71), harbouring the *amdS* gene from *Aspergillus nidulans* as a 2.7 kb *XbaI* fragment, on a *pUC19* vector (Yannisch-Perron et al., 1985. GENE 33:103-119), this plasmid was prepared as described in WO 91/17243.

40 *pDM115*: The preparation is described in Example 1.

pMT1303: The preparation is described in Example 2.

pMT1305: The preparation is described in Example 2.

45 *pMT1329*: The preparation is described in Example 2.

pMT1330: The preparation is described in Example 2.

pMT1332: The preparation is described in Example 2.

pMT1335: The preparation is described in Example 2.

5

Proteolytic Enzyme Assays

Metalloprotease activity is measured as released trypsin activity after a 30-60 min pre-incubation at 25°C in 0.1 M Tris, 2 mM CaCl₂, pH 7 (in the lower pH range a buffer of 100 mM borate, 10 mM DMG, 2 mM CaCl₂ is used). The tryptic activity is measured in
10 microtiter plates; 100 µl samples are mixed with 100 µl of L-BAPNA substrate (a 50-fold dilution in buffer of the stock concentration of 87 mg/ml DMSO, Sigma Aldrich Corp., St. Louis MO, USA) and the absorption at 405 nm is measured using a Thermomax microplate reader from Molecular Devices Corp. (Sunnyvale CA, USA).

The alkaline protease activity is assayed at 25°C using the synthetic substrate, N-succinyl-
15 Ala-Ala-Pro-Phe-p-nitroanilide (Sigma Aldrich Corp.) at 1 mM in 20 mM Tris-HCl (pH 8.0) in a total reaction volume of 1 ml. The release of p-nitroanilin is monitored at 410 nm (Ramesh, MV, et al. 1994. Infection and Immunity 62:79-85).

Example 1

20

Genomic Deletion of the *Aspergillus oryzae* Alkaline Protease (*alp*) Gene and the *Aspergillus oryzae* Neutral Metalloprotease I (*Npl*) Gene

The *alp* and the *Npl* genes were each deleted by the one-step gene replacement method (Miller, B.L., et al., 1985. Mol. and Cell. Biol. 5:1714-1721, and May, G. in Ap-
25 plied Molecular Genetics of Filamentous Fungi, pp. 1-25. J. R. Kinghorn and G. Turner, eds.; Blackie Academic and Professional, 1992.) in two consecutive rounds in a *pyrG*⁻ strain of *A. oryzae*, using the *A. oryzae pyrG* gene as a selection marker.

Cloning of the *Aspergillus oryzae pyrG* gene

30 The *A. oryzae pyrG* gene was cloned by cross-hybridization with the *A. niger pyrG* gene (van Hartingsveldt, W., et al., 1987. Mol. Gen. Genet. 206: 71-75). A lambda phage library from partially Sau3A digested *A. oryzae* IFO 4177 DNA was probed at low stringency with a 1 kb DNA fragment from the *A. niger pyrG* gene. DNA from one positive clone was subcloned into a pUC118 vector. The resultant plasmid, pSO2, was shown to
35 contain the *pyrG* gene by complementation to an *A. niger pyrG* mutant.

Construction of *pyrG* plasmid, pJaL335

A 431 bp fragment located 479 nucleotides upstream from the 5' end of the *A.oryzae pyrG* coding sequence was amplified by PCR, using the following two oligonucleotide primers:

Primer A: 5'-GGAGGAAGATCTCTCTGGTACTCTTCGATCTC -3'; SEQ. ID.

No. 4; and

Primer B: 5'-GGAGGAGAATTCAAGCTTCTTCTACATCACAGTTTGAA
AGC -3'; SEQ. ID. No. 5.

The underlined regions indicate sequences present in the 5' region of the *A. oryzae pyrG* gene. A restriction site was inserted into the 5' end of each primer to facilitate cloning; primer A contains a BglII site and primer B contains an EcoRI site adjacent to a HindIII restriction site.

Plasmid pSO2 was used as template in the PCR reaction. The reaction was performed in a volume of 100 µl containing 2.5 units Taq polymerase, 100 ng of pSO2, 250 nM of each dNTP, and 10 pmol of each of the two primers described above in a reaction buffer of 50 mM KCl, 10 mM Tris-HCl pH 8.0, 1.5 mM MgCl₂.

Amplification was carried out in a Perkin-Elmer Cetus DNA Thermal 480, and consisted of one cycle of 3 minutes at 94°C, followed by 25 cycles of 1 minute at 94°C, 30 seconds at 55°C, and 1 minute at 72°C. The PCR reaction produced a single DNA fragment of 430 bp in length. This fragment was digested with BglII and EcoRI and isolated by gel electrophoresis, purified, cloned into the corresponding site in plasmid pSO2, resulting in a plasmid which was called pJaL335. Thus, pJaL335 harbours the *pyrG* gene flanked by a 431 bp repeat. Fig. 1 summarises the steps in the construction of pJaL335.

Construction of a *pyrG* deletion plasmid, pSO5

A *pyrG* deletion plasmid, pSO5, was constructed from the plasmid pSO2 by digestion with NheI to remove the gene encoding sequence of about 1.1 kb, then religated, leaving about 1 kb each of the *pyrG* 5' and 3' flanking sequences. The construction is illustrated in Fig. 2.

Construction of an *Aspergillus oryzae pyrG* strain, HowB101

A. oryzae IFO 4177 was transformed with the 2.2 kb HindIII fragment from pSO5 harbouring the 5' and the 3' flanking sequence of the *A. oryzae pyrG* gene. The transformants were selected by resistance to 5-fluoro-orotic acid, a phenotype used to select for *pyrG* mutants. One transformant, HowB101, was shown by Southern analysis to contain the expected
5 deletion at the *pyrG* locus. Because HowB101 is uridine dependent, it can be transformed with the wild type *pyrG* gene and transformants selected by the ability to grow in the absence of uridine.

Construction of an *alp* deletion plasmid, pJaL212

10 A genomic DNA library from *A. oryzae* IFO 4177 was obtained by partial digestion with Sau3A. Fragments were ligated into BamHI-digested pUC19. The library was screened using degenerate oligonucleotide probes complementary to known protein sequence fragments of *A. oryzae* alkaline protease. A plasmid containing a 3.9 kb Sau3A fragment was obtained by this screening and named pSRe403.

15 pSRe403 was digested with SacII, then religated, to produce a plasmid in which a fragment of the *alp* gene about 1 kb in length had been deleted. The resultant plasmid was named pSRe403ΔSacII.

pSRe403ΔSacII was partially digested with HindIII, the ends were filled in using Klenow polymerase, then religated to remove the HindIII site in the pUC19 sequence, resulting
20 in a plasmid which was named pJaL173.

A 3.8 kb fragment containing the *pyrG* gene resulting from HindIII digestion of pSO2 was inserted into the HindIII site in pJaL173 to produce a plasmid containing the HindIII fragment of the *pyrG* gene flanked by 5' and 3' sequences of the *alp* gene at the upstream and downstream positions, respectively. Fig. 3 summarizes the construction of plasmid,
25 which was named pJaL212.

Construction of an *Aspergillus oryzae alp*⁻ strain, JaL125

HowB101 was transformed with the 6.8 kb SacI/SphI fragment from pJaL212 using standard procedures. This fragment consists of 1.5 kb of the *alp* promoter, the *pyrG* gene
30 and 1.5 kb of the 3' terminus of the *alp* gene. Transformants were selected by their ability to grow in the absence of uridine.

The transformants were analysed by Southern blotting in which a 599 bp PstI/SacII fragment from pJaL 173 harboring part of the *alp* gene served as a radioactively labelled probe. Strains carrying a deletion of the *alp* gene were recognised by a shift of the wild type PstI band from 1.0 kb to a 1.9 kb band. Four such transformants were identified, one of which was named JaL125.

Cloning of the *Aspergillus oryzae* *Npl* gene

A cosmid library of *A. oryzae* was constructed using the SuperCos1 Cosmid Vector Kit following without modification the instructions provided by the supplier (Stratagene, Inc., La Jolla, CA, USA).

Genomic DNA from *A. oryzae* IFO 4177 was prepared from protoplasts isolated by standard procedures (cf. e.g. Christensen, T., et. al., 1988. Biotechnology 6:1419-1422). After isolation, the protoplasts were pelleted by centrifugation at 2500 rpm for 5 minutes in a Labofuge T (Heto). The pellet was then suspended in a buffer of 10mM NaCl, 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, and treated with 100 µg/ml proteinase K and 0.5% SDS. This and all subsequent steps in the DNA preparation were performed following the recommended procedure of the SuperCos1 Cosmid Vector Kit.

The size of the genomic DNA was analysed by electrophoresis using the CHEF-gel apparatus from BioRad (Hercules, CA, USA). Briefly, a 1% agarose gel was run for 20 hours at 200 volts with a 10-50 second pulse. The gel was stained with ethidium bromide and photographed under UV exposure, revealing DNA ranging from 50 to more than 100 kb in length. The DNA was then partially digested with Sau3A. The size of the resultant DNA fragments ranged from 20 to 50 kb as determined above by CHEF-gel analysis.

The CsCl gradient banded SuperCos1 vector was prepared, ligated and packaged according to the procedures provided with the kit. After titration of the library all of the packaging mix from a single ligation and packaging was transfected into the host cells XL1-Blue MR provided with the vector kit and plated onto LB plates containing 50 µg/ml ampicillin. Approximately 3800 colonies were obtained. Analysis of the cosmid preparation from 10 colonies indicated that all had inserts of the expected size. The colonies were individually picked and inoculated onto 96-well microtiter plates containing 100 µl LB and 100 µg/ml ampicillin per well, then incubated at 37°C overnight. One hundred microliters of 50% glycerol was added to each well, and the entire library was frozen at -80°C. A total of 3822

colonies were stored, representing an approximate 4.4 fold amplification of the *A. oryzae* genome.

Preparation of a *Fusarium oxysporum* p45 metalloprotease probe

5 The *Fusarium oxysporum* p45 metalloprotease gene was cloned as disclosed in patent application WO 95/30757 (Novo Nordisk Biotech, Inc.), published 16 November 1995.

A clone from the above cDNA library was selected and designated pDM115. A probe was prepared from pDM115 containing a 1.76 kb fragment of the *Fusarium oxysporum* cDNA that encodes a portion of the p45 gene. This plasmid was digested with Sall, and the
10 fragments were separated on a 1% agarose gel. The band of interest was a 1.5 kb fragment from which the DNA was eluted. This fragment was labelled with ³²P-dATP by random priming and used as a probe for Southern analysis and to screen an *A. oryzae* cosmid library.

Screening of the *A. oryzae* library

15 To screen the *A. oryzae* cosmid library with the *Fusarium oxysporum* p45 probe, the individually frozen colonies in the library were inoculated onto LB plates containing 100µg/ml ampicillin by using a multipin device in a configuration of 6 rows of 8 pins, designed to fit into half of a 96-well microtiter dish. Plates, containing colonies from all clones in the library, were incubated at 37°C overnight. Sterilized Whatman 540 filters cut
20 to fit the petri dishes were placed over the colonies which were then incubated for two more hours at 37°C. The filters were transferred to LB plates containing 200µg/ml of chloramphenicol, and the plates were incubated overnight at 37°C. The next day the filters were washed twice in 0.5M NaOH for 5 minutes, then twice in 0.5M Tris-HCl (pH 7.4) for 5 minutes and finally twice in 2x SSC for 5 minutes. The filters were wet down with ethanol
25 and allowed to air dry.

The filters were hybridised with the 1.5 kb ³²P-labelled DNA fragment from pDM115. The hybridisation was carried out for 16 hours at 65°C in a standard hybridisation buffer of 10x Denhart's solution, 5x SSC, 0.02 M EDTA, 1% SDS, 0.15 mg/ml polyA RNA and 0.05 mg/ml yeast tRNA. After hybridisation the filters were washed in 2x SSC, 0.1% SDS
30 at 65°C twice and exposed to X-ray films. Three colonies showed hybridisation to the probe, 3E8, 3C1 and 2A5.

Characterisation of the cosmid clones

Restriction analysis revealed that two of the three cosmid clones, 3E8 and 3C1, contain inserts derived from the same region of the *A. oryzae* genome. Three micrograms of cosmid DNA was digested with EcoRI and fractionated by agarose gel electrophoresis. The DNA
5 was transferred to Immobilan-N membrane filters (Millipore) and hybridised with the radio-labelled pDM115 probe. The results revealed that the probe hybridised to a 4 kb EcoRI fragment in both cosmid clones. The 4.0 kb EcoRI fragment was subjected to further analysis.

10 Construction of the *Npl* deletion plasmid, pJaL399

Plasmid pToC65, described in EP 531 372, was digested with SacI, and treated with bacterial alkaline phosphatase according to the manufacturer's instructions (Boehringer Mannheim) to remove the 5' phosphate groups, then phenol extracted and precipitated. The
15 5.5 kb SacI fragment from cosmid clone 3E8 containing the *A. oryzae Npl* gene was isolated by gel electrophoresis and purified. The two fragments were mixed together and ligated. After transformation into *E.coli*, the colonies carrying the correct plasmid were identified by restriction enzyme digestion of mini-plasmid preparations; the recovered plasmid was called pJaL389. The presence of the coding region of the *A. oryzae Npl* gene was confirmed by DNA sequence analysis of portions of this subclone.

20 Plasmid pJaL389 was digested with Ball to delete an internal 1.1 kb fragment of the *Npl* gene. The remaining 7.1 kb fragment was treated with Klenow polymerase to create blunt ends, then isolated by gel electrophoresis and purified. This DNA fragment was next treated with bacterial alkaline phosphatase, phenol extracted and precipitated. Plasmid pJaL335 was digested with HindIII to obtain a 3.5 kb fragment comprising the *A.oryzae*
25 *pyrG* gene, then similarly treated with Klenow polymerase, isolated by gel electrophoresis, and purified. The two fragments were mixed together and ligated. After transformation into
* *E.coli*, the colonies carrying the correct plasmid were identified by restriction enzyme digestion of mini-plasmid preparations. The construction of this plasmid, called pJaL399, is summarised in Fig. 4.

30 Thus, plasmid pJaL399 contains a fragment of the pToC65 vector harbouring the *Npl* gene in which the internal 1.1 kb Ball fragment has been replaced with a 3.5 kb DNA fragment encoding the *A. oryzae pyrG* gene.

Isolation of a *pyrG*⁻ *A. oryzae* strain, JaL151

The *A. oryzae alp*⁻ strain, JaL125, was screened for resistance to 5-flouro-orotic acid to identify spontaneous *pyrG* mutants. One strain, named JaL151, was identified as being *alp*⁻ and *pyrG*⁻. As with HowB101, JaL151 is uridine dependent, therefore it can be transformed with the wild type *pyrG* gene and transformants selected by the ability to grow in the absence of uridine.

Construction of an *Aspergillus oryzae NpI* strain, JaL228

JaL151 was transformed with the 7.9 kb *SacI* fragment of pJaL399 using standard procedures. Transformants were then selected by their ability to grow in the absence of uridine. After reisolation chromosomal DNA was prepared from 12 transformants. The DNA from each of the transformants was digested with *BalI* and analysed by Southern blotting, using the 5.5 kb ³²P-labelled DNA *SacI* fragment from pJaL389 containing the *NpI* gene as described above as the probe. Strains of interest were identified by the absence of the 1.1 kb *BalI* hybridising fragment, indicating a disruption of the *NpI* gene, as well as by showing an altered mobility of the flanking 5' and 3' sequences. The strain resulting from one of these transformants was named JaL228.

20 Example 2

Cloning of *C. antarctica* lipase B

A partial *Sau3A* digested *Candida antarctica* genomic library in pBR322 was prepared and replica filters were screened with two oligonucleotide probes designed from the N-terminal amino acid sequence of *clb*, NOR929 (SEQ. ID. No. 4) and NOR930 (SEQ. ID. No. 5) (Fig. 5). Eight colonies were identified. Analysis of the plasmids indicated overlapping inserts. One of the plasmids, pMT1303, containing a 7.8 bp insert, was further analysed. By deleting the *SalI* fragment a plasmid, pMT1305, containing a 2.1 kb insert was constructed (Fig. 6). Using NOR929 as a template, pMT1305 was sequenced to confirm that *clb* (*C. antarctica* lipase B) was encoded. The entire gene was sequenced in both strands by making a series of deletions in pMT1305 using the Promega Erase-a-Base© system (Promega Corp., Madison WI, USA). The coding sequence is given in SEQ. ID. No. 6.

Construction of a partial CLB plasmid, pMT1329

A 788 bp SalI-HindIII fragment encoding the N-terminal sequence for CLB was isolated from pMT1305. From this fragment two smaller fragments were isolated, a 170 bp HindIII-BanI and a 204 bp BanI filled-in Sau3A. These two fragments were cloned into the plasmid pIC19R, which had been digested with EcoRV and HindIII. This created a BclI site at the EcoRV and filled-in Sau3A site 9 bp adjacent to the ATG codon for *clb*, resulting in the plasmid named pMT1329 (Fig. 7).

10 Construction of a partial CLB plasmid, pMT1332

The gene sequence encoding the C-terminal region of CLB was isolated as a 670 bp HindIII-XmnI fragment from pMT1305 and inserted into plasmid pIC7, which had been digested with HindIII-NruI, resulting in pMT1330. The complete gene was assembled from the 3.0 kb HindIII-NarI fragment from pMT1329 and the 0.7 kb HindIII-ClaI fragment from pMT1330, resulting in pMT1332 (Fig. 8).

Construction of a CLB expression plasmid, pMT1335

The *clb* gene was expressed in *Aspergillus* using the expression vector pToC68. The 1 kb BclI-BglII fragment from pMT1332 was inserted into pToC68, which had been digested with BamHI, to give pMT1335 (Fig. 9). Thus, the construction of the plasmid pMT1335 resulted in a fungal expression plasmid for the *C. antarctica* lipase B gene.

Expression of *C. antarctica* lipase B in *Aspergillus oryzae* strains

A. oryzae strains IFO 4177 and JaL228 were transformed with pMT1335 by co-transformation with pToC90 (need to define pToC90).

Transformants were selected for growth on minimal medium containing 10 mM acetamide and screened for the presence of pMT1335 by the ability to produce CLB. One transformant from each of the *A. oryzae* wild type strain IFO 4177 and the *alp*⁻, *NpI* strain JaL228 was selected for tank fermentation run for 93 hours. Conidiospores were inoculated into 2-liter Keiler fermentors at 30°C containing 1.2 litre 20% maltose liquor and 8% urea. A mixture of 20% maltose liquor and 8% urea was added continuously; phosphoric acid was added as necessary to maintain a pH of 7.

Summarised in Table 1 below is the yield of CLB activity from JaL228 in comparison to IFO4177, in which one lipase unit (LU) is defined as the amount of enzyme needed to liberate 1 μ mol of butyric acid per minute from tributyrin at pH 7.0 and at 30°C. The table indicates that the amount produced in JaL228 is 1.8 times greater than in IFO 4177, demonstrating that the deletion of the *alp* and *Npl* genes can significantly improve the stability of proteins sensitive to the activity of these proteases.

Hours	IFO4177 LU/g	JaL228 LU/g
93	575	1041

Table 1. CLB production by IFO 4177 and JaL228 after tank fermentation for 93 hours.

10

15

SEQUENCE LISTINGS

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Novo Nordisk A/S
- (B) STREET: Novo Alle
- (C) CITY: Bagsvaerd
- (E) COUNTRY: Denmark
- (F) POSTAL CODE (ZIP): DK-2880
- (G) TELEPHONE: +45 4444 8888
- (H) TELEFAX: +45 4449 3256

(ii) TITLE OF INVENTION: Novel Host Cells and Methods of Producing Proteins

(iii) NUMBER OF SEQUENCES: 6

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30B (EPO)

(2) INFORMATION FOR SEQ. ID. No. 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2052 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Fusarium oxysporum*
- (B) STRAIN: DSM 2672
- (C) INDIVIDUAL ISOLATE: p45

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 785..2049

(ix) FEATURE:

- (A) NAME/KEY: sig_peptide
- (B) LOCATION: 55..784

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 364..415

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 802..854

(ix) FEATURE:

(A) NAME/KEY: intron

(B) LOCATION: 1821..1868

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 1:

ATGCGTTTCT CCGACTCTCT CCTCCTCATC GGCCTATCCA GCCTCGCTGG TGCTCATCCC	60
AGCAGAAGGG CTCCTAATCC TTCACCGCTG AGCAAGCGTG GCCTCGACCT GGAAGCTTTT	120
AAGCTTCCTC CCATGGCCGA GTACGTTCTT CAGGACGAGG TTCCTGATGA TGTCAGTGCC	180
AAGGTCGTCA CCAAGCGCGC TGATTACACC GAGACTGCCA AGGACTTGGT TAAGTCGACT	240
TTCCCCAAGG CTACTTTCCG TATGGTCACG GATCACTATG TTGGTAGCAA CGGAATTGCG	300
CATGTAAACT TTAAGCAGAC TGTCACGGT ATTGATATCG ACAATGCTGA TTTCAACGTC	360
AACGTGGGTA TTCTCAAGAC TTTGGGGAGT TTGGAATGTG CTGACATGGA TACAGATTGG	420
CGCTGACGGC GAGGTCTTCT CCTACGGAAA CAGCTTCTAC GAGGGCAAGA TTCCCGGTCC	480
TCTTACCAAG CGTGACGAGA AAGACCCCGT CGACGCTCTC AAGGACACCG TTGATGTTCT	540
TTCTCTCCCC GTTGAGGCTG ACAAGGCCAA GGCTGAGAAG AAGAGCAAGA ACCACTACAC	600
CTTCACTGGT ACCAAGGGTA CCGTCAGCAA GCCCGAGGCT AAGCTCACCT ACCTTGTTGA	660
TGAGAACAAG GAGCTCAAGC TCACATGGAG AGTTGAGACT GATATTGTTG ACAACTGGCT	720
GTTGACTTAT GTCAATGCTG CCAAGACTGA TGAGGTTGTT GGTGTTGTTG ACTACGTCAA	780
TGAGGCGACA TACAAGGTCT AGTACGTATT TCCATAAATT GACGATTGGG AAAGAATTGA	840
CCGTTGTATT ATAGTCCTTG GGGTGTCAAT GATCCCTCCA AGGGATCTCG CTCCACTGTT	900
GAGAACCCCT GGAATCTCGC GGCCTCCGAG TTCACCTGGC TCAGCGACGG CTCAAACAAC	960
TACACCACAA CCCGCGGGAA CAATGGAATT GCACAGGTGA ATCCTTCAGG GGGCTCCACG	1020
TATCTGAACA ATTACCGTCC TGATAGCCCG TCGCTGAAGT TCGAGTATGA TTACTIONACC	1080
AGCACCACCTA CACCCACCAC CTACCGCGAT GCTTCCATCG CTCAGCTTTT CTACACAGCC	1140
AACAAGTACC ACGACCTCCT CTACCTTCTT GGCTTTACCG AACAGGCTGG TAACTTCCAG	1200
ACCAACAACA ATGGCCAGGG TGGTGTAGGA AACGATATGG TTATCCTCAA CGCTCAGGAC	1260
GGAAGCGSCA CCAACAACGC CAACTTCGCT ACACCCGCTG ACGGTCAGCC CGGCCGCATG	1320
CGAATGTATC TCTGGACATA CAGCACACCC CAGCGTGAAT GCAGTTTCGA CGCTGGCGTT	1380
GTTATCCACG AGTACACTCA CGGTCTCTCC AACCCTCTCA CAGGTGGCCC TGCCAACTCG	1440
GGTTGTCTTC CCGGTGGTGA ATCCGGTGGC ATGGGTGAGG GCTGGGGTGA CTTTCATGGCT	1500
ACTGCCATTC ACATCCAATC CAAGGATACC CGCGCTAGCA ACAAGGTCAT GGGTGACTGG	1560
GTGTACAACA ACGCAGCTGG TATCCGAGCT TATCCTTACA GTACAAGCCT TACCACTAAC	1620
CCTTACACTT ACAAGAGTGT TAACAGTCTC AGTGGAGTCC ATGCTATTGG TACTTACTGG	1680

```

GCTACTGTTC TGTATGAGGT TATGTGGAAC CTCATCGACA AGCATGGGAA GAATGATGCG      1740
GATGAGCCCA AATTCAACAA CGGCGTTCCT ACAGATGGCA AATATCTTGC TATGAAGTTA      1800
GTAGTGGATG GCATGTCGCT GTAAGTTGTC CCTTGGATTT GTAGGAGTTC TTATCTAACG      1860
TTTAATAGGC AACCTTGCAA CCCCAACATG GTCCAGGCCC GAGACGCCAT CATCGACGCC      1920
GACACCGCTC TTACCAAGGG AGCTAACAAG TGCAGATCT GGAAGGGCTT TGCCAAGCGT      1980
GGTCTTGGA CTGGTGCCAA GTATAGTGCT TCCAGCCGTA CTGAGAGCTT TGCTCTTCCT      2040
TCTGGATGTT AA                                                                2052

```

(3) INFORMATION FOR SEQ. ID. No. 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2968 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus oryzae*
- (B) STRAIN: IFO 4177
- (C) INDIVIDUAL ISOLATE: NpI

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 458..817

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 868..1262

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 1320..1870

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 1930..2344

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 2404..2587

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 2:

CCCCCGATTG GGAAGACCTT GAATCCAATC TATCGGAGAG CCGACGGAAC GGGTTTTCTT

GAACCCAATG TCCCAATTGC TTTGAGAAAA TGGATATTAT TCAATGCCCC TGATCATTAT	120
TGGCATACTC TTCACGGAGT GGTCTGCCCC TGGGCGATTG CTCCACCATG AGGTGCATGA	180
AGGTAACCCC GAGCTGCTAT AGATGAAGTG TGTCGGATAT TGCAGGAGTC TGAATGCGCA	240
GCTGATTGCC CTTTTCCCAC CATGATGAAT GGGACACAAT ATGTCTAGCT CCGGGTTGTC	300
TCGGGCCAAG AAGAAATGCT TCTTTGCGTT GATCCAGCGC GTCGAAATCT CCAATTGAAC	360
CGGTATAAAT AACAGCTGAG CGTCGCAGCA TTCAAGGCAG AACCATTCCA CAGATCATCT	420
TCTATCATTT GCATTGAGTC CAACGCCCTG GCAGTTCATG ATGCGGGGTC TTCTACTAGC	480
TGGAGCCCTT GGCCTACCTT TGGCCGTCCT TGCGCATCCG ACCCATCATG CACATGGACT	540
TCAACGTCGC ACAGTTGACT TGAATCATT CCGTTTGCAC CAGGCAGCGA AGTATATCAA	600
TGCGACTGAG TCTTCGAGTG ATGTTTCATC TTCTTTCTCT CCCTTCACCG AGCAAAGCTA	660
CGTGGAGACG GCCACTCAGC TCGTGAAGAA TATCCTGCCA GATGCTACCT TCCGTGTCGT	720
CAAGGATCAT TACATTGGTA GCAATGGGGT CGCTCATGTC AATTTTCGTC AGACGGTCCA	780
TGGCCTTGAC ATTGACAATG CGGACTTCAA TGTCAATGTA CGCTGCAGTC CACCTATACT	840
ATGTTCCGGT CTAACCACTT CATTTAGGTT GGGAAAAATG GAAAGATCTT TTCCTATGGC	900
CACTCATTTT ATACGGGCAA AATCCCCGAT GCCAATCCTT TGACGAAGCG GGATTATACC	960
GACCCTGTAG CGGCTCTCAG AGGAACCAAC GAAGCTTTAC AGCTTTCTAT CACTCTAGAT	1020
CAAGTGTCTA CTGAGGCTAC CGAGGACAAA GAGTCCTTCA ATTTCAAGGG AGTCTCTGGC	1080
ACCGTTTCGG ATCCCAAGGC TCAGTTGGTC TACTTGGTAA AGGAAGATGG CAGCCTTGCT	1140
TTGACCTGGA AGGTGGAGAC AGATATTGAC AGCAACTGGC TGTTGACCTA CATCGATGCC	1200
AATACCGGCA AAGATGTCCA TGGTGTGGTT GACTACGTAG CCGAGGCAGA TTACCAAGTA	1260
TAGTGAGTAT TTTAAGAATG TGACTTGGAC TGTAGAATGA AGCTGACACA CCACCACAGT	1320
GCATGGGGTA TTAATGATCC CACGGAGGGC CCTCGCACCG TCATCAGCGA TCCATGGGAT	1380
TCGTCCGCAT CTGCGTTCAC CTGGATCAGT GACGGAGAAA ACAACTATAC CAACTCGC	1440
GGCAACAACG GTATCGCGCA GTCGAACCT ACCGGTGGAT CGCAGTACTT GAAGAACTAC	1500
CGGCCTGATA GCCCCGATTT GAAATTCCAA TACCCCTATT CGCTCAACGC CACACCCCCA	1560
GAGTCCTATA TTGATGCGTC TATCACTCAG CTTTTCTACA CTGCCAACAC GTACCACGAC	1620
CTACTCTACA CTCTGGGCTT CAACGAGGAG GCCGGTAATT TCCAGTACGA TAACAATGGA	1680
AAAGGAGGTG CTGGAAACGA CTACGTGATC CTCAATGCTC AGGACGGTTC TGGCACCAAT	1740
AACGCCAACT TCGCTACGCC CCCGGATGGA CAGCCCGGCC GCATGCGCAT GTACATTTGG	1800
ACCGAGTCCC AGCCTTACCG TGACGGCTCC TTCGAGGCTG GTATTGTGAT TCACGAGTAT	1860
ACTCACGGTC GTATGTATCC CTTATGAACC CCAAGATAAG GCAGTCTGAA CTAACACCAT	1920
GGTACACAGT CTCTAACCGG CTCACTGGAG GACCTGCTAA CTCTCGCTGC TTGAATGCCC	1980
TTGAATCCGG CGGAATGGGT GAAGGTTGGG GAGACTTCAT GGCCACGGCA ATTCGGCTCA	2040
AGGCCGGCGA TACTCACTCG ACCGATTATA CCATGGGTGA ATGGGCTGCA AACAAGAAAG	2100

```

GTGGCATCCG TGCTTACCCA TTCTCAACCT CCCTGGAAAC CAACCCCTCTC ACCTACACCA      2160
GTCTCAATGA ATTGGACGAA GTGCATGCCA TCGGCGCGGT GTGGGCTAAC GTATTGTACG      2220
AGCTGTTGTG GAACTTGATC GATAAGCACG GCAAGAATGA CGGGCCAAAG CCCGAGTTCA      2280
AGGATGGAGT TCCGACTGAC GGCAAGTATC TCGCCATGAA GCTGGTGATT GATGGCATAG      2340
CATTGTAAGT GCCAACCTCG TTTCCTCTTT CTACCTATCG CAGGGGCTAA CCTTGACTTT      2400
TAGGCAACCT TGCAACCCCA ACTGTGTCCA GGCTCGCGAC GCCATCCTCG ATGCCGATAA      2460
GGCTCTCACC GATGGTGCTA ACAAGTGCGA GATTTGGAAG GCGTTTGCTA AGCGTGGTTT      2520
GGGTGAAGGC GCTGAATACC ATGCGTCTCG TCGGGTGGGC AGTGATAAGG TGCCCTCTGA      2580
TGCTTGCTAG AGTTTGTGTA TATTCTTTCA GCTAGGTGAT TGGGGAGATC CTGTAGGCTG      2640
ACTATAGTTT GAATTTAACA ACATTTCTTA ATTCTTACTG TCAGGTAACT TGCCTGCAC      2700
AATAGCATGA ATAGGTTATG GCACTCTGGC CCTTATGTCG GATTCTTACA CAATGATTTG      2760
TCCTTAAGAA ATACTATCTC AACTCATCAC CCCTCTCATC GTCCGCCGCT TCGAATTCAG      2820
GGTCATTCTT CCCCCTCTTC CTGGCCCCCT TGGCCGTCCT GACCCTCTGC TTGCTCTCCA      2880
TCTCCCCCTC CACCCTGATC ATAGGTCACG GTGGCCGCCG GTCCCACCAG GTGGTTGGGG      2940
CAATAGTTGT TTATAGCACG GGGCTGCT      2968

```

(4) INFORMATION FOR SEQ. ID. No. 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3922 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Aspergillus oryzae*
- (B) STRAIN: A01560
- (C) INDIVIDUAL ISOLATE: alp

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 2310..2634

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 2684..3129

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3188..3276

- (ix) FEATURE:
(A) NAME/KEY: exon
(B) LOCATION: 3333..3686

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 3:

GGATCCATAA TGAGCACACT TGAACCTCGC ATGAGTGCTC CATATCTTAG ATCCATTGTA	60
GTCACGAGCG TCGAAGCCAA GGCTAAAAGA AAAGGGACCA AGACCGCCGC ATATCGGGGT	120
GGCTGAACAA TCGCTGCTAG TGTGTACACA GGATACATAA ATTAATCAAG AGCCAATGAC	180
GTTATCGAAG ACATGATATT ACTTACTTGA TTTGTTGCGG TGCTCCGACT CAACTGGATT	240
CAGATGGAAG GGAAGGGATT CTTGATCAGG CTCGTCCCGG GATGACCATT GCTCCTCTCG	300
AATCGAGACA ATGTTGAATG CCCAGGATGA CGATTGAGAG CGAGATATTA AAAAGAGAT	360
TAAGGGTTGA CGATTGCCAA CGGAAGCCCG ATGGAAGAAG AAATGCCAAG AATTTGGCTG	420
CCTGGATGTT AGTGCCGTTT AGCCTCAGGC CCCAGCTGAA TCCCTGCCAA TCACGAGGCA	480
GGGCTCACCA CCTCCAACCC TTACACAAGA CTCGCCCTCG CTCTTTCCTG CAGGTCCTCT	540
GCTTACTTTC CCTCTTTTC CCTCCTGAAA TGCCCCGAGA ATGCCGTCCG GAGCTTAGGA	600
AAGCTACGAC GGCTTAATGT TCTAATTTC CCCACCGACC TGCCCTGGCC AGTTAGACCG	660
CCGGACCCAT CGAATGCAAC ATCACCAACA CTAATTTACC TGCATCTCTG TCAGCCCACT	720
GGGTTTAACT AGATATGCCA AGGACTTGCT TGGCTGGTTT ATGCATGAAG AGAGATGGGC	780
ACTAGTGCGT GCGGGACCAC AAACCTCACC TGCAGAGGGC TCATGCACCT GAAAGACTGC	840
CAATGATCAT TTGACTGGTT AGGTCAAGGG GTTAGGCTTA GAGCCCTTTG CTAATGCCGA	900
TGCCGCCTCT TTGACTGCCA CATTTCCTGG CTTCCTCTCT TGGCCCTCC CGTCCCTTGA	960
TGCCAAGGGC CTTGGTGGCT CGGACTCCCG GCGGTAGGCT GGCCACCTAC TCATGGCGTG	1020
CGGAGATGGG CCATTCAGGT TGTGCCTAAT GACATACTCT GATGCCGACG GGAAGGCCGG	1080
CTGCTTGCTG GTGTCATTGG CTTCTCGAAT GACTCGGAGG ATTGTCGTCT TGCAGGACTT	1140
TTGTGTAACA ACAGGGGCCG ACAGTTGGCG TGGCTGCCGT GGATATGCTT TTECTGCGAG	1200
CCATGGTTAT TCTGCGGAAC GAAACCACCC TCCCACCCAA ACAGGGCTAA TGTGCCCAGG	1260
TCCTGATACC ATCAGAAGAC CTCCAGGAGC ACATGCCTGT TCGCATAACC GTGGTGTAGC	1320
ACCAGGAATT GCTTAGCTTA GCTTCTTCGA CTGGGGGGCC AGAAAGTGCT TATCGCAAAG	1380
ATCCCACTTC TTTGTGTGAT AGCCCCTCCC GCGGCCCTTG ATCAAGCCGT TCTCGCTCGC	1440
CCATACCGAA ACCGCGATAT TATAGGTGCA GATGGTTATT ATTCTTTTTT TTTTCTTTT	1500
TCTTTGCTTC TCATGCAGCC CCATACGTTG CCGAATTTGG CTACACCTTG GGGCTCATTC	1560
TTCGAAGTTT AGATTCCGAC AAGACCTCAG CACCCAATCA AAACCCTTGA TTCCTGATAA	1620
AAGACGTGGA AAAAAGCGGA TATCGCGTGA GGATGCCAAG CAAAGGGAAT GGGTCACATT	1680
GATCTCTGTC GCGCTGTTAG GATGATCTTC ACTCCTAAAG GCATCGCCGC GGCATTAGGC	1740

CCTTCCCTGT CCAAGATATC GGTACTCCT CTCATTATGG CGAGCTACTT TGTGAATTAA	1800
TTGACTGAGG GATATAACCAC CTTCCCTTTG AAGGTACCGA GCCACTACCT TGAGCGTTAG	1860
TTACTTTTTT GAGGAAAGCA TCCTATGCTA GTCTCTGCCA ATCACTGCAG CGTCGACAAC	1920
TTGCCATAGC CTTGTGTTCT TCACGGTCTA TCGGAACACC CGTTCATGAC TGAAAGGGGT	1980
CAGCGTCCGT GGTGGTCAAC ATCATTCTCA TCTTTTCATCA TGCCCGCTGA TTGATAGAGT	2040
AATTTCCGGT GGAGCACAAAC GCCGTCCTCT GAGATGCAAT GTCACCCTGT AAGTTTCAAC	2100
TACAATCTGT AGTACAGAGC ATCCTTGCTAC TTGCATGCTG TGCAAGTGAT CCAAATCCGT	2160
AGAACTTGCT CGAGAACAGG GAAATATAGA ACTCCTGAAG GTTATAAATA CCACATGCAT	2220
CCCTCGTCCA TCCTCACTTC CATCATCAAG CCAGCGGTTT CTATCCTCCG ACTTGAGTTG	2280
TTCTTGCGCA TCTTTACAAT CTTCTCATCA TGCAGTCCAT CAAGCGTACC TTGCTCCTCC	2340
TCGGAGCTAT CCTTCCCGCG GTCCTCGGTG CCCCTGTGCA GGAAACCCGC CGGGCCGCTG	2400
AGAAGCTTCC TGGAAAGTAC ATTGTCACAT TCAAGCCCGG CATTGACGAG GCAAAGATTC	2460
AGGAGCATAC CACCTGGGCT ACCAACATTC ACCAGCGCAG TCTGGAGCGT CGTGGCGCCA	2520
CTGGCGGTGA TCTTCCTGTC GGTATTGAGC GCAACTACAA GATCAACAAG TTCGCCGCTT	2580
ATGCAGGCTC TTTGACGAT GCTACCATTG AGGAGATTCT CAAGAACGAA GATGTTTGTG	2640
GTCATCCGCT CGCATTTTTG AATGACAGCT AACTCGCGCC CAGGTTCCTT ACGTCGAGGA	2700
GGACCAGATC TACTACCTCG ATGGCCTGAC TACCCAGAAG AGTGCCCCCT GGGGTCTGGG	2760
CAGCATTTCC CACAAGGGCC AGCAGAGCAC CGACTACATC TACGACACTA GTGCCGGCGA	2820
GGGCACCTAT GCCTACGTGG TGGATAGCGG TGTCAATGTC GACCATGAGG AGTTCGAGGG	2880
CCGCGCCAGC AAGGCCTACA ACGCTGCCGG TGGTCAGCAT GTGGACAGCA TTGGCCATGG	2940
CACCCACGTT TCCGGCACCA TTGCTGGCAA GACTTATGGT ATCGCCAAGA AGGCCAGCAT	3000
CCTTTCGGTC AAAGTTTTCC AGGGTGAATC GAGCAGCACT TCCGTCATTC TTGACGGCTT	3060
CAACTGGGCT GCCAACGACA TTGTTAGCAA GAAGCGTACC AGCAAGGCTG CAATCAACAT	3120
GAGCTTGGGT GAGTTTACAT TGTCTTCTC TACTTGGAAC GCGCGAGCGC TAATTTCAA	3180
AACACAGGCG GTGGCTACTC TAAGGCTTTC AACGATGCGG TCGAGAACGC ATTGAGCAG	3240
GGTGTCTCT CGGTTGTCGC TGCCGGTAAC GAGAACGTAC GTCTCCCTC CATCGCGCAA	3300
AGACGAATTC GTAACGACT TGATTTCTT AGTCTGATGC CGGCCAAACC AGCCCTGCCT	3360
CTGCCCCGTA TGCCATCACT GTTGCCGCTA TCCAGAAGAG CAACAACCGC GCCAGTTTCT	3420
CCAACCTTGG CAAGGTCGTT GACGCTTTCG CTCCCGGTCA AGATATCCTT TCTGCCTGGA	3480
TTGGCTCTTC CTCTGCCACC AACACCATCT CTGGTACCTC CATGGCTACT CCCACATTG	3540
TCGGCCTGTC CCTCTACCTC GCTGCCCTTG AGAACCTCGA TGGCCCCGCT GCCGTGACCA	3600
AGCGCATCAA GGAGTTGGCC ACCAAGGACG TCGTCAAGGA TGTTAAGGGC AGCCCTAACC	3660
TGCTTGCTTA CAACGGTAAC GCTTAAGTAC CAGGAGTACG TCGCAGGATT CTACCATTGT	3720
TACTGGAATA CAATGATGAT TAGAAAACGA AGAGCGTTAT GATTGCGACG GATATATGCA	3780

TGGCACCCAT ACAGCGTGAT ACATAGGCTG TTGCTCAAG AATTAGGATT TTATCTGAAT 3840
CCATGTACAG AGTATACTTA TGTTAGTAGT CAATAAAATC TTGGCTTTCT AATTTTGTCC 3900
CATCTACAAG GGGTCGTCGA TC 3922

(5) INFORMATION FOR SEQ. ID. No. 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 4:

CCCTCGGGCT CGGACCCCGC CTTCTCGCA GCCCAAG

(6) INFORMATION FOR SEQ. ID. No. 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 5:

CCCTCGTTTG TCAGGCCCGC GTCCAGCACC GACTTGGGCT G

(7) INFORMATION FOR SEQ. ID. No. 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1029 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Candida antarctica*

(C) INDIVIDUAL ISOLATE: clb

(xi) SEQUENCE DESCRIPTION: SEQ. ID. No. 6:

ATGAAGCTAC TCTCTCTGAC CGGTGTGGCT GGTGTGCTTG CGACTTGCCT TGCAGCCACT	60
CCTTTGGTGA AGCGTCTACC TTCCGGTTTCG GACCCTGCCT TTTCGCAGCC CAAGTCGGTG	120
CTCGATGCGG GTCTGACCTG CCAGGGTGCT TCGCCATCCT CGGTCTCCAA ACCCATCCTT	180
CTCGTCCCCG GAACCGGCAC CACAGGTCCA CAGTCGTTTCG ACTCGAACTG GATCCCCCTC	240
TCAACGCAGT TGGGTTACAC ACCCTGCTGG ATCTCACCCC CGCCGTTTCAT GCTCAACGAC	300
ACCCAGGTCA ACACGGAGTA CATGGTCAAC GCCATCACCG CGCTCTACGC TGGTTTCGGGC	360
AACAACAAGC TTCCCGTGCT TACCTGGTCC CAGGGTGGTC TGGTTGCACA GTGGGGTCTG	420
ACCTTCTTCC CCAGTATCAG GTCCAAGGTC GATCGACTTA TGGCCTTTGC GCCCGACTAC	480
AAGGGCACCG TCCTCGCCGG CCCTCTCGAT GCACTCGCGG TTAGTGCACC CTCCGTATGG	540
CAGCAAACCA CCGGTTCCGG ACTCACCACC GCACTCCGAA ACGCAGGTGG TCTGACCCAG	600
ATCGTGCCCA CCACCAACCT CTA CTCTCGGCG ACCGACGAGA TCGTTCAGCC TCAGGTGTCC	660
AACTCGCCAC TCGACTCATC CTACCTCTTC AACGGAAAGA ACGTCCAGGC ACAGGCCGTG	720
TGTGGGCCGC TGTTCGTCAT CGACCATGCA GGCTCGCTCA CCTCGCAGTT CTCCTACGTC	780
GTCGGTCGAT CCGCCCTGCG CTCCACCACG GGCCAGGCTC GTAGTGCAGA CTATGGCATT	840
ACGGACTGCA ACCCTCTTCC CGCCAATGAT CTGACTCCCG AGCAAAAGGT CGCCGCGGCT	900
GCGCTCCTGG CGCCGGCAGC TGCAGCCATC GTGGCGGGTC CAAAGCAGAA CTGCGAGCCC	960
GACCTCATGC CCTACGCCCG CCCCTTTGCA GTAGGCAAAA GGACCTGCTC CGGCATCGTC	1020
ACCCCTGA	1029

CLAIMS

1. A host cell useful for the expression of a heterologous protein product, in which said cell has been genetically modified in order to express significantly reduced levels of activity of a metalloprotease and an alkaline protease activity, as compared to a parental cell.
2. The host cell according to claim 1, which is a fungal cell.
3. The host cell according to claim 1, which is a yeast cell.
4. The host cell according to claim 3, which is a strain of *Saccharomyces*, preferably *Saccharomyces cerevisiae*.
5. The host cell according to claim 2, which is a filamentous fungus.
6. The host cell according to claim 5, which is a strain selected from the group consisting of *Acremonium*, *Aspergillus*, *Candida*, *Cocliobolus*, *Endothia*, *Fusarium*, *Humicola*, *Neurospora*, *Rhizomucor*, *Rhizopus*, *Thermomyces*, *Trichoderma*, *Podospora*, *Pyricularia*, and *Penicillium*.
7. The host cell according to claim 6, which is a strain selected from the group consisting of *Aspergillus nidulans*, *Aspergillus awamori*, *Aspergillus phoenicis*, *Aspergillus japonicus*, *Aspergillus foetus*, *Fusarium oxysporum*, *Fusarium solani*, *Humicola grisea*, *Neurospora crassa*, *Penicillium chrysogenum*, *Rhizomucor meihei*, *Trichoderma reesei*, and *Trichoderma viride*.
8. The host cell according to claim 6, which is a strain selected from the group consisting of *Aspergillus oryzae*, *Aspergillus niger* and *Fusarium* species having the identifying characteristics of ATCC 20334.
9. The host cell according to any of claims 1-8, in which the metalloprotease is a *Fusarium* metalloprotease.

10. The host cell according to claim 9, in which the metalloprotease is a *Fusarium oxysporum* metalloprotease.
11. The host cell according to claim 10, in which the metalloprotease is a *Fusarium oxysporum* p45 metalloprotease having the nucleotide sequence presented as SEQ. ID. No. 1, or a sequence homologous thereto.
12. The host cell according to any of claims 1-11, in which the metalloprotease is a neutral metalloprotease, having optimal proteolytic activity in the range of about pH 6-8.
- 10 13. The host cell according to claim 12, in which the metalloprotease is a neutral *Aspergillus* metalloprotease of group Npl or NpII.
14. The host cell according to claim 13, in which the metalloprotease is an *Aspergillus oryzae* neutral metalloprotease I preferably one comprising an amino acid sequence derived from the nucleotide sequence presented as SEQ. ID. No. 2, or a sequence homologous thereto.
- 15 15. The host cell according to any of claims 1-14, in which the alkaline protease is an *Aspergillus* alkaline protease.
- 20 16. The host cell according to claim 15, in which the alkaline protease is preferably one encoded by a cDNA sequence comprising the nucleotide sequence presented as SEQ. ID. No. 3, or a sequence homologous thereto.
- 25 17. The host cell according to any of claims 1-16, which has been genetically modified at the structural or regulatory regions encoded in the metalloprotease and the alkaline protease genes.
18. The host cell according to claim 17, which has been genetically modified by specific or random mutagenesis, PCR generated mutagenesis, site specific DNA deletion, insertion and/or substitution, gene disruption or gene replacement techniques, anti-sense techniques, or a combination thereof.
- 30

19. The host cell according to any of claims 1-18, in which the level of metalloprotease is reduced more than about 50%, preferably more than about 85%, more preferred more than about 90%, most preferred more than about 95%.
- 5 20. The host cell according to any of claims 1-18, in which the level of alkaline protease activity is reduced more than about 50%, preferably more than about 85%, more preferred more than about 90%, most preferred more than about 95%.
21. The host cell according to any of claims 1-18, in which the reduction of metalloprotease
10 and alkaline protease activity is any combination according to claim 19 or 20.
22. The host cell according to any of claims 1-21, in which the cell is essentially free of any metalloprotease and alkaline protease activity.
- 15 23. A method of producing a host cell according to any one of claims 1-22 wherein a parental cell is modified to express significantly reduced levels of activity of a metalloprotease and an alkaline protease, as compared to a parental cell.
24. A method of producing a host cell according to any of claims 1-22 wherein said cell has
20 been genetically modified within structural and/or regulatory regions encoded in the metalloprotease and alkaline protease genes.
25. A method of producing a heterologous protein product in the host cell of any one of claims 1-22, in which the method comprises:
- 25 (a) introducing into said host cell a nucleic acid sequence encoding said protein product;
- (b) cultivating in a suitable growth medium, the host cell of step (a); and
- (c) isolating said heterologous protein product.
- 30 26. The method according to claim 25, in which the protein product is a therapeutically active gene product, such as insulin, growth hormone, glucagon, somatostatin, interferon, EPO, TPO,

PDGF, factor VII, factor VIII, urokinase, chymosin, tissue plasminogen activator, or serum albumin.

27. The method according to claim 25, in which the protein product is a protein of fungal
5 origin.

28. The method according to claim 27, in which the protein product is a fungal enzyme, in particular an amylolytic enzyme, such as an α -amylase, a β -amylase, a glucoamylase, a β -galactosidase, a cellulytic enzyme, a lipolytic enzyme, a xylanolytic enzyme, a proteolytic
10 enzyme, an oxidoreductase, such as a peroxidase or a laccase, a pectinase, or a cutinase.

29. The method according to any of claim 25, in which the protein product is a bacterial protein.

15 30. The method according to any of claims 25-29, in which the protein product is a precursor protein, i.e. a zymogen, a hybrid protein, a protein obtained as a pro sequence or pre-pro sequence, or in an immature form.

1/9

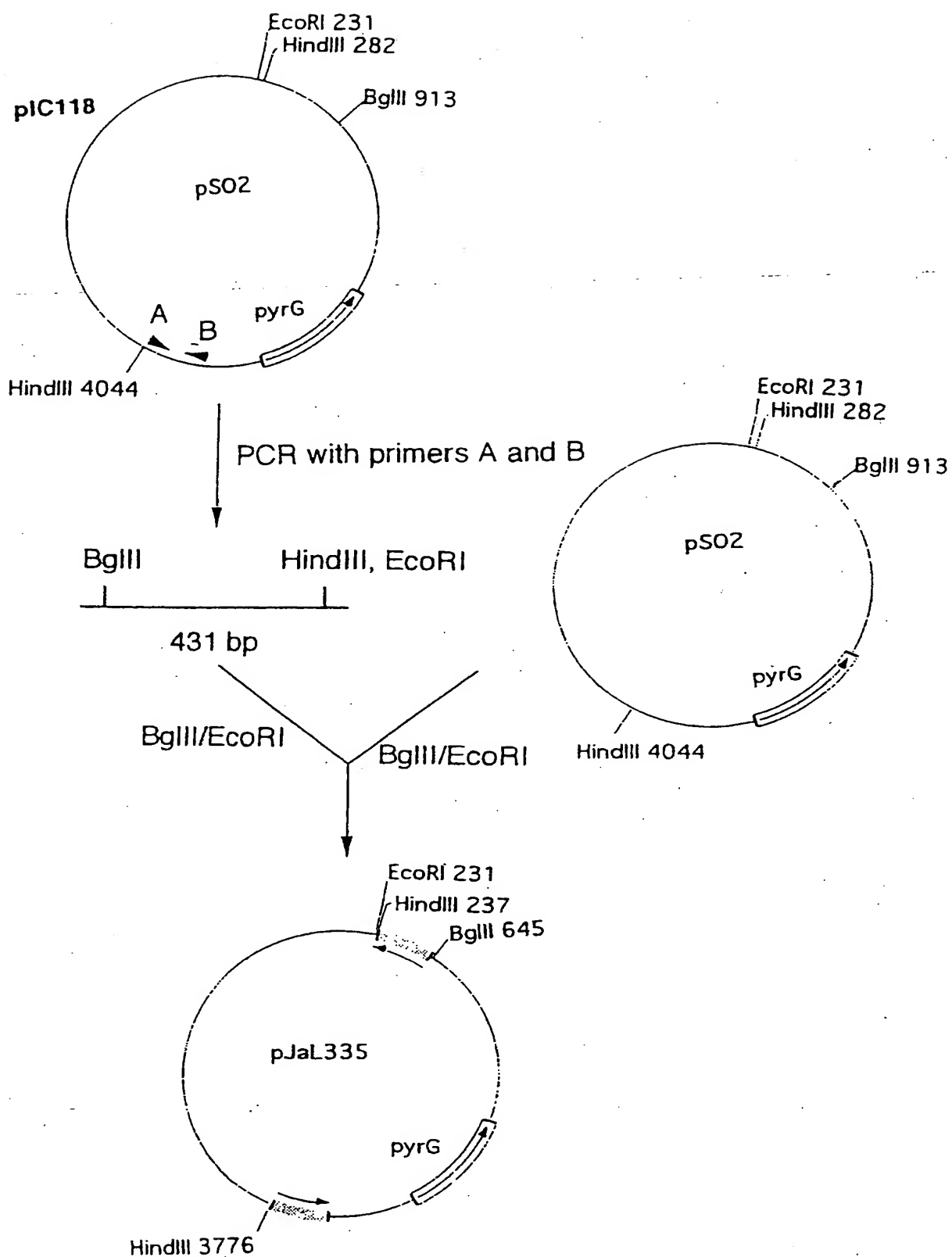


Fig. 1

2/9

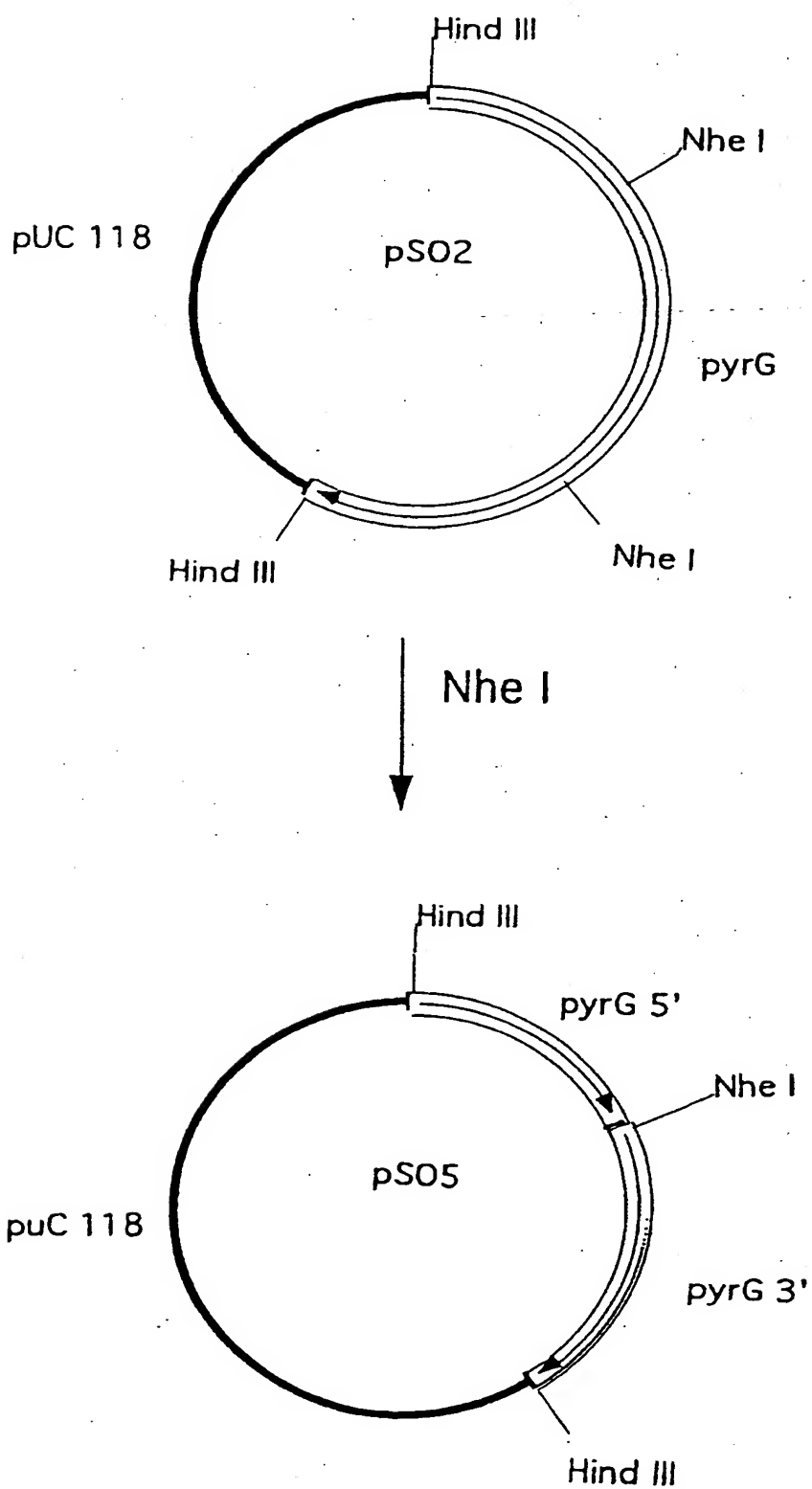


Fig. 2

3/9

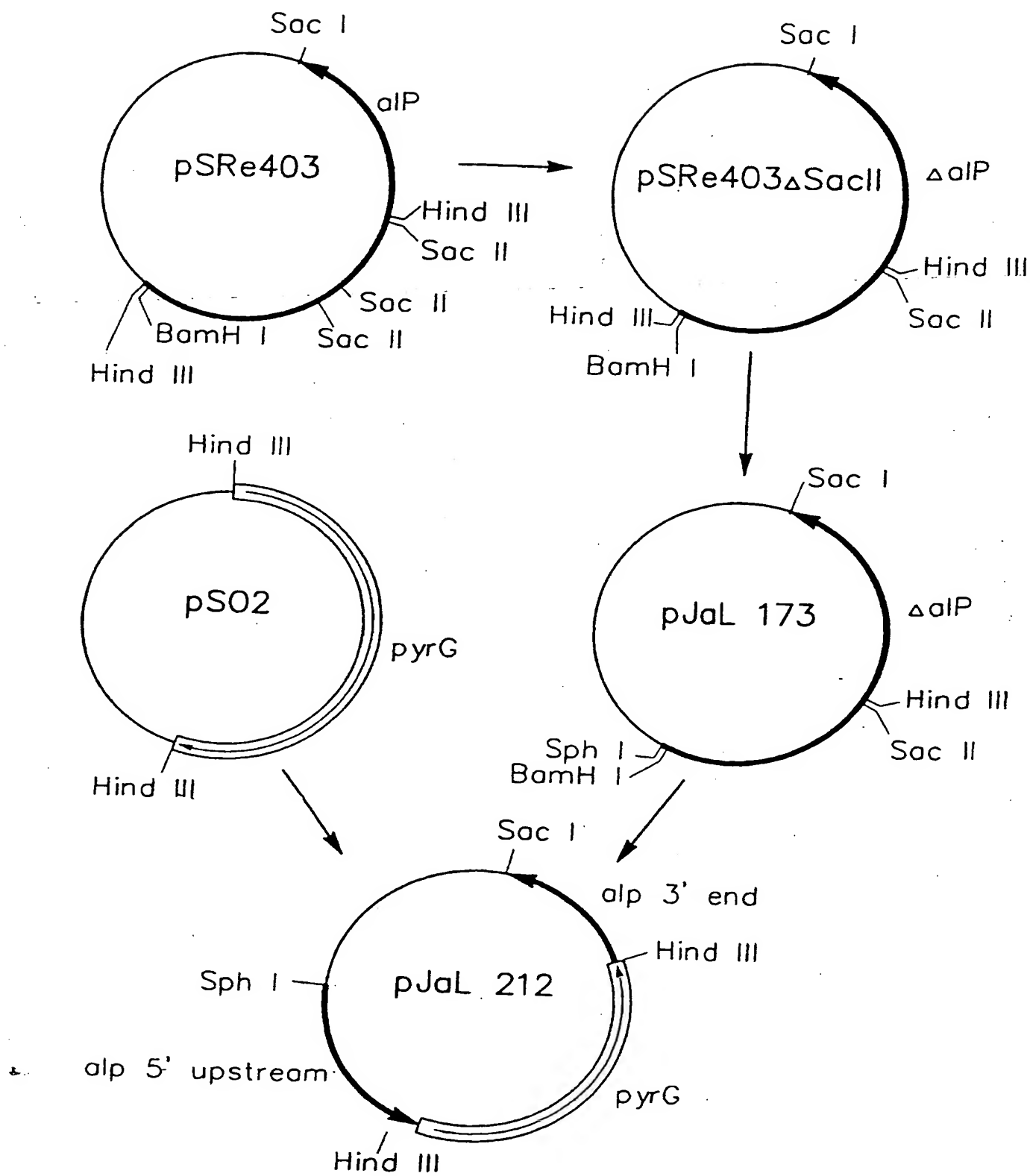


Fig. 3

4/9

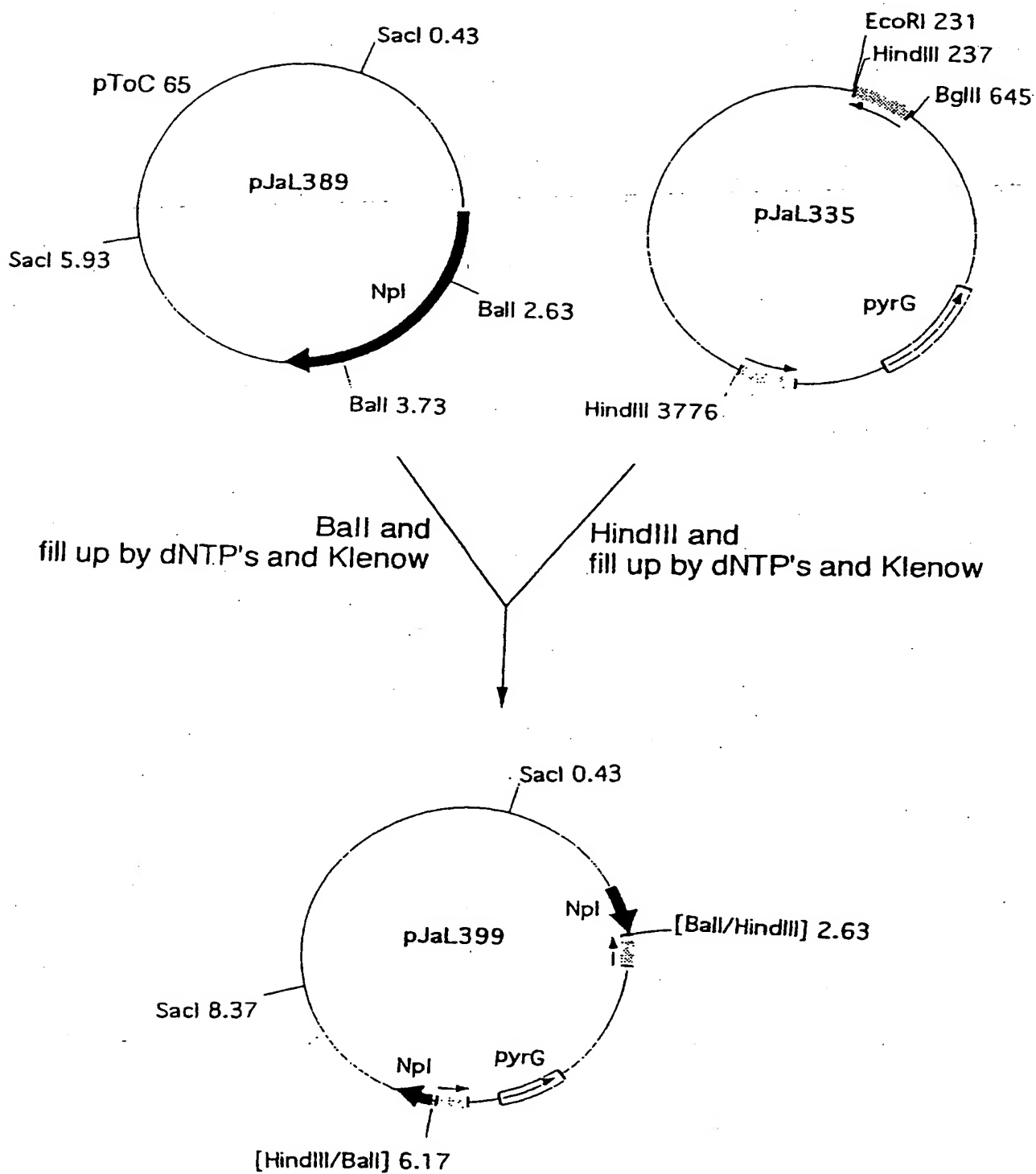
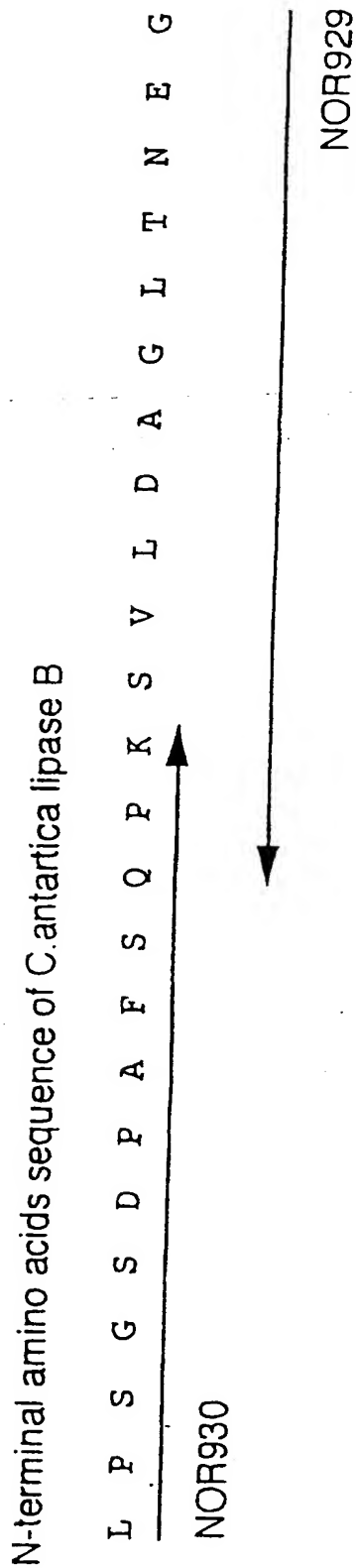


Fig. 4

5/9



NOR929

CCCTCGGGCTCGGACCCCGCCTTCTCGCAGCCCAAG

G G G T

NOR930

CCCTCGTTCGTCAGGCCCGGTCAGCACCAGCTTGGGCTG

G G G G

Fig. 5

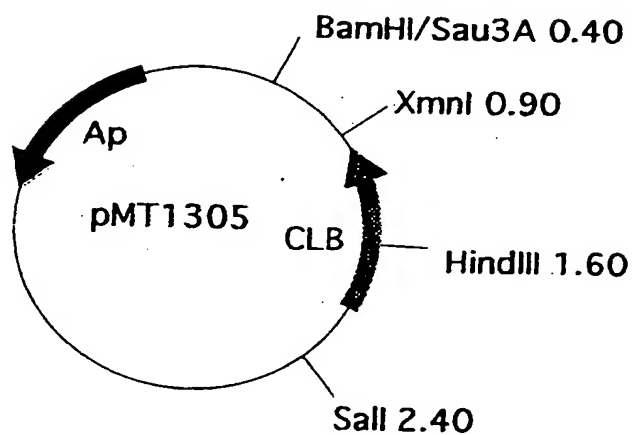
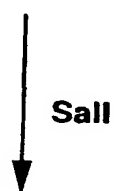
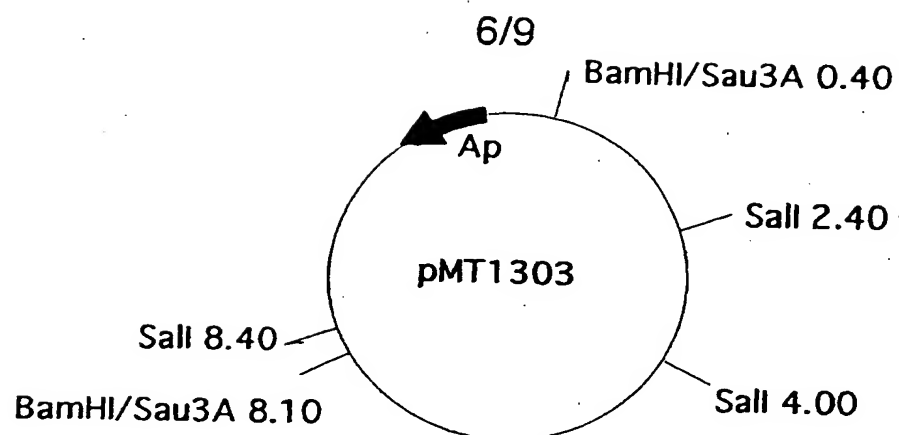


Fig. 6

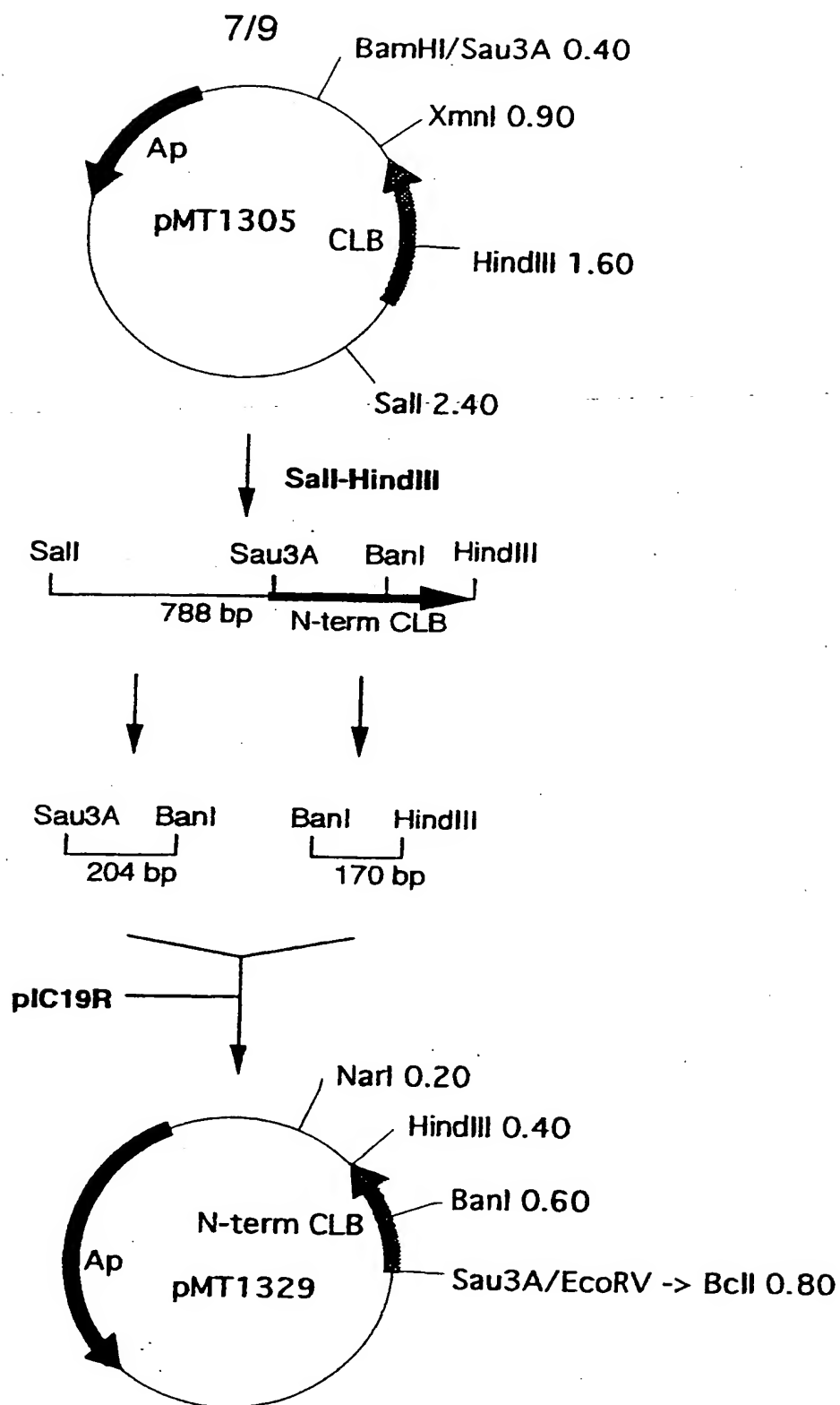


Fig. 7

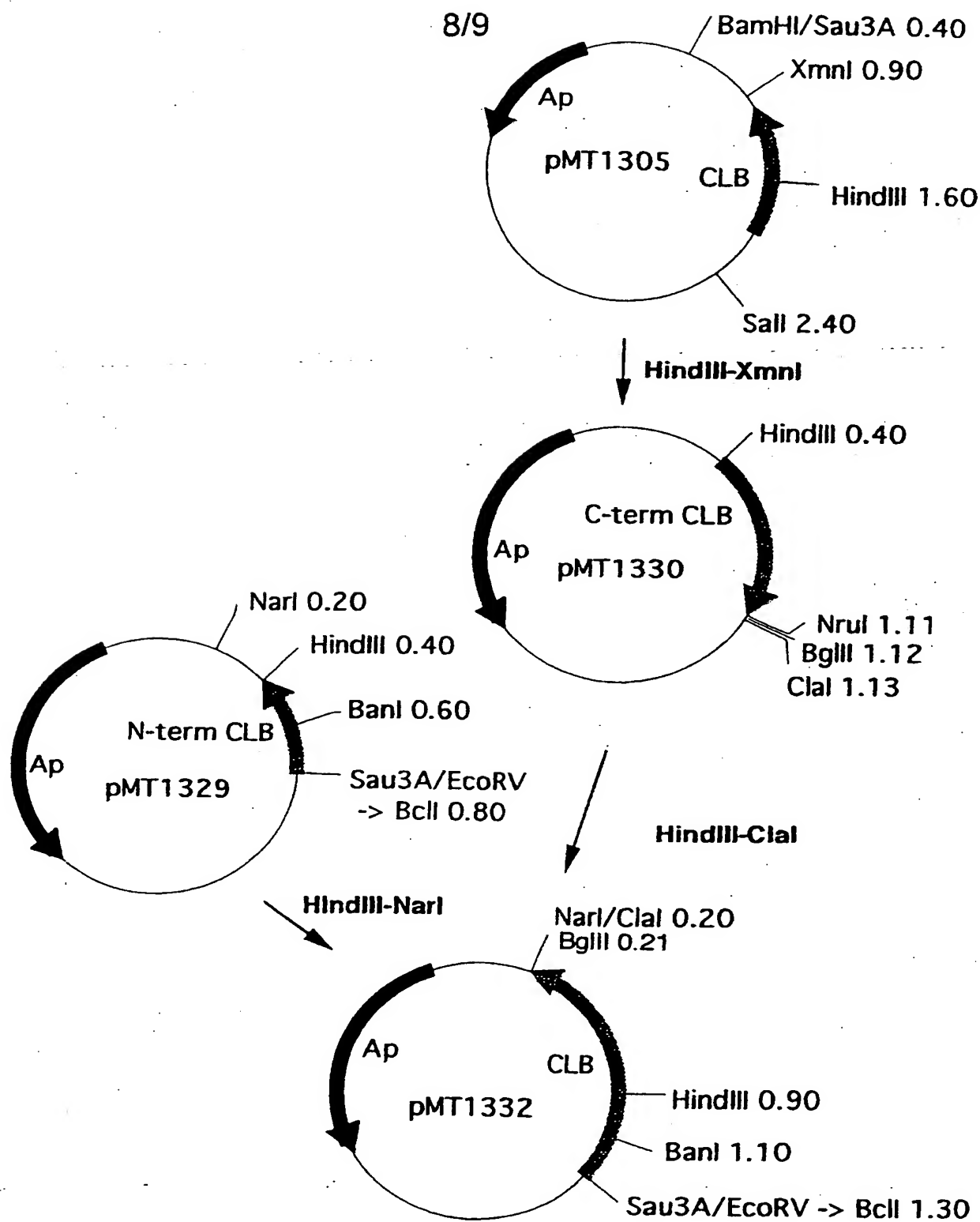
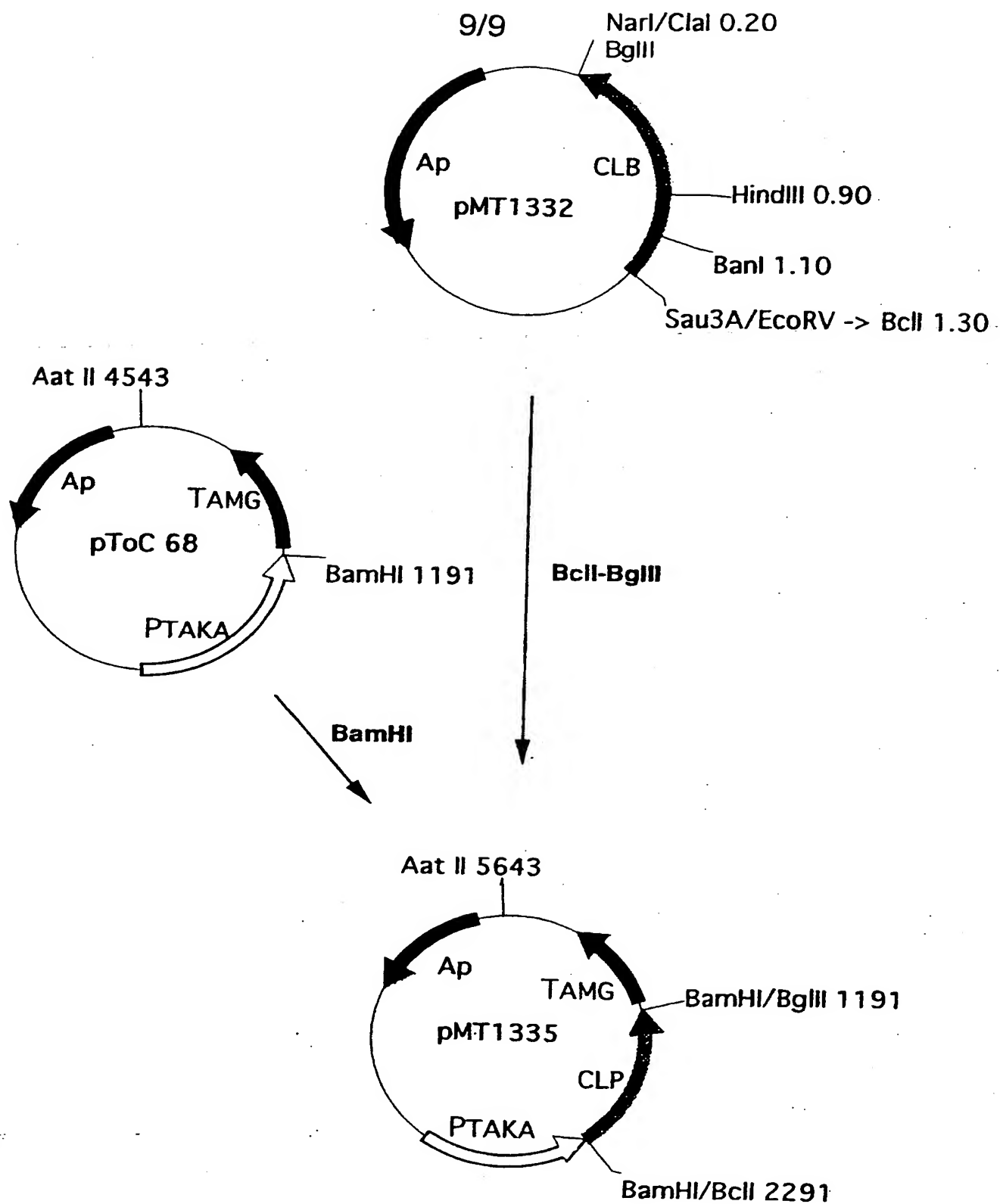


Fig. 8

**Fig. 9**

SUBSTITUTE SHEET (RULE 26)

International application No.

PCT/DK 97/00397

A. CLASSIFICATION OF SUBJECT MATTER

IPC6: C12N 1/19, C12N 1/15, C12N 15/80, C12N 15/81
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC6: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, MEDLINE, BIOSIS, EMBASE, EMBL/SWISSPROT/PIR/GENESEQ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	WO 9629391 A1 (NOVO NORDISK A/S), 26 Sept 1996 (26.09.96), see page 6, lines 13-14 and the claims --	1-30
X	WO 8904866 A1 (CELLTECH LIMITED), 1 June 1989 (01.06.89), see page 1, lines 27-33; page 2, lines 20-22; page 17, line 12-14 and the claims --	1-30
X	WO 9217595 A1 (THE SALK INSTITUTE BIOTECHNOLOGY/INDUSTRIAL ASSOCIATES, INC.), 15 October 1992 (15.10.92), page 1, line 24 - page 2, line 3; page 7, line 11 --	1-30

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

11 December 1997

09-01-1998

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. +46 8 666 02 86

Authorized officer

Carolina Palmcrantz
Telephone No. +46 8 782 25 00

SUBSTITUTE SHEET (RULE 26)

International application No.

PCT/DK 97/00397

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0327797 A1 (BOEHRINGER MANNHEIM GMBH), 16 August 1989 (16.08.89), see the claims --	1-30
X	Molecular Microbiology, Volume 14, No 5, 1994, Katia Jatou-Ogay et al, "Cloning and disruption of the gene encoding an extracellular metalloprotease of Aspergillus fumigatus" page 917 - page 928 -----	1-30

SUBSTITUTE SHEET (RULE 26)

International application No.
PCT/DK 97/00397

Patent document cited in search report			Publication date	Patent family member(s)		Publication date
WO	9629391	A1	26/09/96	AU	4940096 A	08/10/96
WO	8904866	A1	01/06/89	AU	2727388 A	14/06/89
				DK	367689 A	27/09/89
				EP	0351427 A	24/01/90
				JP	2502247 T	26/07/90
WO	9217595	A1	15/10/92	AU	661844 B	10/08/95
				AU	1750592 A	02/11/92
				AU	3660095 A	11/01/96
				CA	2105064 A	02/10/92
				EP	0578746 A	19/01/94
				JP	6506117 T	14/07/94
				US	5324660 A	28/06/94
				US	5541112 A	30/07/96
EP	0327797	A1	16/08/89	DE	3804890 A,C	13/07/89
				JP	2002385 A	08/01/90
				US	5179003 A	12/01/93